BJPsych Open (2017)
3, 265–273. doi: 10.1192/bjpo.bp.117.005058

Individual differences in schizophrenia

Edmund T. Rolls, Wenlian Lu, Lin Wan, Hao Yan, Chuanyue Wang, Fude Yang, Yunlong Tan, Lingjiang Li, Chinese Schizophrenia Collaboration Group, Hao Yu, Peter F. Liddle, Lena Palaniyappan, Dai Zhang, Weihua Yue and Jianfeng Feng

Background
Whether there are distinct subtypes of schizophrenia is an important issue to advance understanding and treatment of schizophrenia.

Aims
To understand and treat individuals with schizophrenia, the aim was to advance understanding of differences between individuals, whether there are discrete subtypes, and how first-episode patients (FEP) may differ from multiple episode patients (MEP).

Method
These issues were analysed in 687 FEP and 1880 MEP with schizophrenia using the Positive and Negative Syndrome Scale for (PANSS) schizophrenia before and after antipsychotic medication for 6 weeks.

Results
The seven Negative Symptoms were correlated with each other and with P2 (conceptual disorganisation), G13 (disturbance of volition), and G7 (motor retardation). The main difference between individuals was in the cluster of seven negative symptoms, which had a continuous unimodal distribution. Medication decreased the PANSS scores for all the symptoms, which were similar in the FEP and MEP groups.

Conclusions
The negative symptoms are a major source of individual differences, and there are potential implications for treatment.

Declaration of interests
L.P. received speaker fees from Otsuka Canada and educational grant from Janssen Canada in 2017.

Previous research has indicated that schizophrenia is heterogeneous in clinical symptoms and time course, suggesting the possibility of identifying subtypes differing in cause and/or pathological mechanism.[1] The identification of subtypes could facilitate investigation of cause and mechanism, including the use of genetic and neuroscience approaches.[2–5] Many approaches consider that there are positive symptoms (including hallucinations and delusions), negative symptoms (including low mood, motivation, and emotion), and cognitive symptoms, including poor attention.[6–8] It has been suggested that at least in chronic cases, the positive symptoms split into two subgroups: reality distortion and disorganisation.[7,9] Many studies that have included a wider range of symptoms have found evidence for five clusters of symptoms, namely, reality distortion, disorganisation, negative symptoms, depression and excitation (reviewed by Peralta & Cuesta).[10]

In order to better understand and treat individuals with schizophrenia, the aims of the present investigation were to understand differences between individuals with schizophrenia; whether there are discrete subtypes of schizophrenia; how first-episode patients (FEP) may differ from multiple episode patients (MEP) and how medication influences the symptoms. Two approaches to delineating the heterogeneity of schizophrenia were taken. One approach was an examination of the pattern of correlation between symptoms to identify underlying symptom communities (or dimensions), presumed to reflect specific pathological processes. Another approach was an examination of the correlations between patients, to identify communities of patients (subtypes) each characterised by a symptom profile reflecting a particular combination of symptom dimensions. In the main text, we employ the standard and well-established hierarchical clustering method using k-means.[11] In the supplement, we present qualitatively similar results obtained using a powerful community detection algorithm based on the concept of modularity described by Newman & Girvan,[12] which was also useful in providing an estimate of the optimum number of communities.

Method

Participants and clinical assessment

Two samples comprising a total of 2567 patients with schizophrenia were diagnosed according to DSM-IV criteria.[13] The first sample included 687 FEP and drug-naive patients with schizophrenia. The second study sample included 1880 multi-episode patients (MEP) with schizophrenia with average duration of illness 7.97 years. The patients were recruited from in-patients from four research centres (the Sixth Hospital of Peking University, the Second Xiangya Hospital of Central South University, Beijing Anding Hospital and Beijing HuiLongGuan Hospital) with institutional approval of the investigation and informed consent of the patients.

The symptoms were assessed by trained researchers with the 30-item Positive and Negative Syndrome Scale for schizophrenia (PANSS).[14] For reference, in PANSS, positive symptoms P1–P7 are Items 1–7, the negative symptoms N1–N7 are Items 8–14 and the general symptoms G1–G16 are Items 15–30. In the figures, unless otherwise stated, the symptoms are presented in this order, 1–30, with the symptoms listed in the legend to Fig. 1a. The PANSS was measured first in each of these patients when they were not receiving drug treatment. Then the patients were provided with one out of the seven randomly assigned drugs (risperidone, quetiapine, perphenazine, olanzapine, haloperidol, aripiprazole and ziprasidone) for a 6-week treatment period, and then the PANSS scores were measured again. This design with the same patients assessed

before and after medication allowed measurement of the PANSS when untreated, and also enabled the effects of the medication to be assessed using exactly the same patients. Further details of the patient populations and the medication used are provided in the Data supplement.

Statistical analysis

The data consisted for the MEP group of 1880 patient×30 symptom matrix, and for the FEP group of a 687 patient×30 symptom matrix, for the unmedicated pre-treatment state. Corresponding matrices were available for the post-treatment state, providing the symptoms in the same patients after 6 weeks of treatment. The methods included calculating the symptom correlation matrix and the population correlation matrix. We also examined whether there were discrete populations of patients using the k-means clustering method (performed on the population correlation matrix), and then analysed the symptoms in the sub-populations that were detected. We also investigated the distributions of symptoms across the patient sub-populations.

The data were analysed using the k-means cluster analysis algorithms that are a well-accepted approach standardised for availability within Matlab, using the distance measure based on correlations. This analysis was complemented by a modern community detection method described in the Data supplement, which has the advantage that it can automatically provide an indication of the optimal number of subgroups or communities into which to cluster the data.[15,16] The k-means approach was the method used for analysis described in the main text, because it is a well-known and understood method for clustering data, and because one major difference between the populations that was identified with this approach was in the negative symptoms which have a component that has a unimodal continuous distribution, so we did not wish to rely on a community detection method that might have difficulty if it was forced to determine the number of communities into which to divide the population if it contained at least in part a continuous distribution. However, for comparison key results with the community detection method are provided in the Data supplement.

Results

Multi-episode group pre-treatment (MEP_Pre)

Figure 1a shows the MEP_Pre symptom correlation matrix. The negative symptoms (8–14) form a clear cluster. Some correlation between these negative symptoms and P2 (conceptual disorganisation), G11 (poor attention), G13 (disturbance of volition) and G16 (active social avoidance) are also evident in the symptom correlation matrix shown in Fig. 1a. The community detection method indicated that the symptoms belong to four distinguishable communities (Fig. S1). Symptom community four included the majority of the positive symptoms (P1 P4 P5 P6 P7) and general symptoms (G8 G9 G12 G14). Symptom Community 1 included the majority of the negative symptoms (N1 N2 N3 N4 N6) and general symptoms (G7 G13 G16).

As shown in Fig. S2, the community detection algorithm indicated that the sample contained three distinguishable communities of patients. The patient clusters generate by the three-cluster output of the k-mean algorithm had distinct symptom profiles (Fig. 1b). In the MEP_Pre sample, the sub-population labelled PN (positive and high negative symptoms, 694 cases) had high values for the negative symptoms (N1–N7=symptoms 8–14). The sub-population labelled Pn (positive and intermediate negative symptoms, 535 cases) had intermediate values for the negative symptoms. The sub-population labelled P (positive and low negative symptoms, 651 cases) had low values for the negative symptoms, especially N1 and N2. Little else differed between these clusters, except that the PN cluster had a higher value for P2 (conceptual disorganisation) and relatively small values for P1 and P3–P6; whereas the P cluster had relatively high scores on symptoms P1, P3–P7, G8 and G9.

These analyses provided evidence that the scores of the negative symptoms were an important factor that was different between the three populations detected with k-means, and equally by the community detection algorithm (see Data supplement Fig. S3). Furthermore, Fig. 1a shows that there are moderately high correlations between the seven negative symptoms, N1–N7 (Symptoms 8–14). To investigate how the negative symptoms were distributed across the population of MEP patients, the population × symptom data matrix was sorted by the mean value of symptoms N1–N7. The result shown in Fig. 2a indicates that there is a graded distribution of the negative symptoms across the population of 1880 MEP patients, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible). (In a control analysis, it was found that this does not apply to the positive symptoms, which are not continuously graded throughout the range between scores of 1 and 7 as are the negative symptoms shown in Fig. 2a.)

Further evidence was then sought of whether the distribution of the mean score for the seven negative symptoms was continuous and unimodal. Figure 2b and c show that the distribution of the mean of the average of the negative symptom scores in the MEP_Pre group has a unimodal distribution, in one dimension, rather than a bimodal or multimodal distribution. Thus the most important way in which the three sub-populations of MEP–Pre patients differed was in the negative symptom scores, which are continuously distributed from 1 to 7.

Multi-episode group post-treatment (MEP_Post)

The MEP_Post symptom correlation matrix showed, similarly to the MEP_Pre condition illustrated in Fig. 1a, that the negative symptoms (8–14) form a clear cluster (Fig. S5). Three sub-population clusters were again found by the community detection algorithm and the three clusters in the MEP_Post dataset that were detected with the k-means algorithm which had 631 patients in the PN cluster, 643 in the Pn cluster and 606 in the P cluster.

Figure 1c shows the average symptom values in the three patient clusters for the MEP_Post group. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms (N1–N7=symptoms 8–14). The group labelled Pn (positive and intermediate negative symptoms) had intermediate values for the negative symptoms. The group labelled P (positive and low negative symptoms) had low values for the negative symptoms and relatively low values also for the positive symptoms. It is notable that the scores on all or most of the symptoms have been decreased by the medication (Fig. 4d). Just as in the MEP_Pre condition, a continuously graded distribution of the negative symptoms across the population of 1880 MEP_Post patients was found, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible).

First-episode group pre-treatment (FEP_Pre)

Figure 3a shows the FEP_Pre symptom correlation matrix. The negative symptoms (8–14) form a clear cluster.

The community detection algorithm detected three clusters or sub-populations of patients, and Fig. 3b shows the average symptom values in the three patient clusters for the FEP_Pre group from

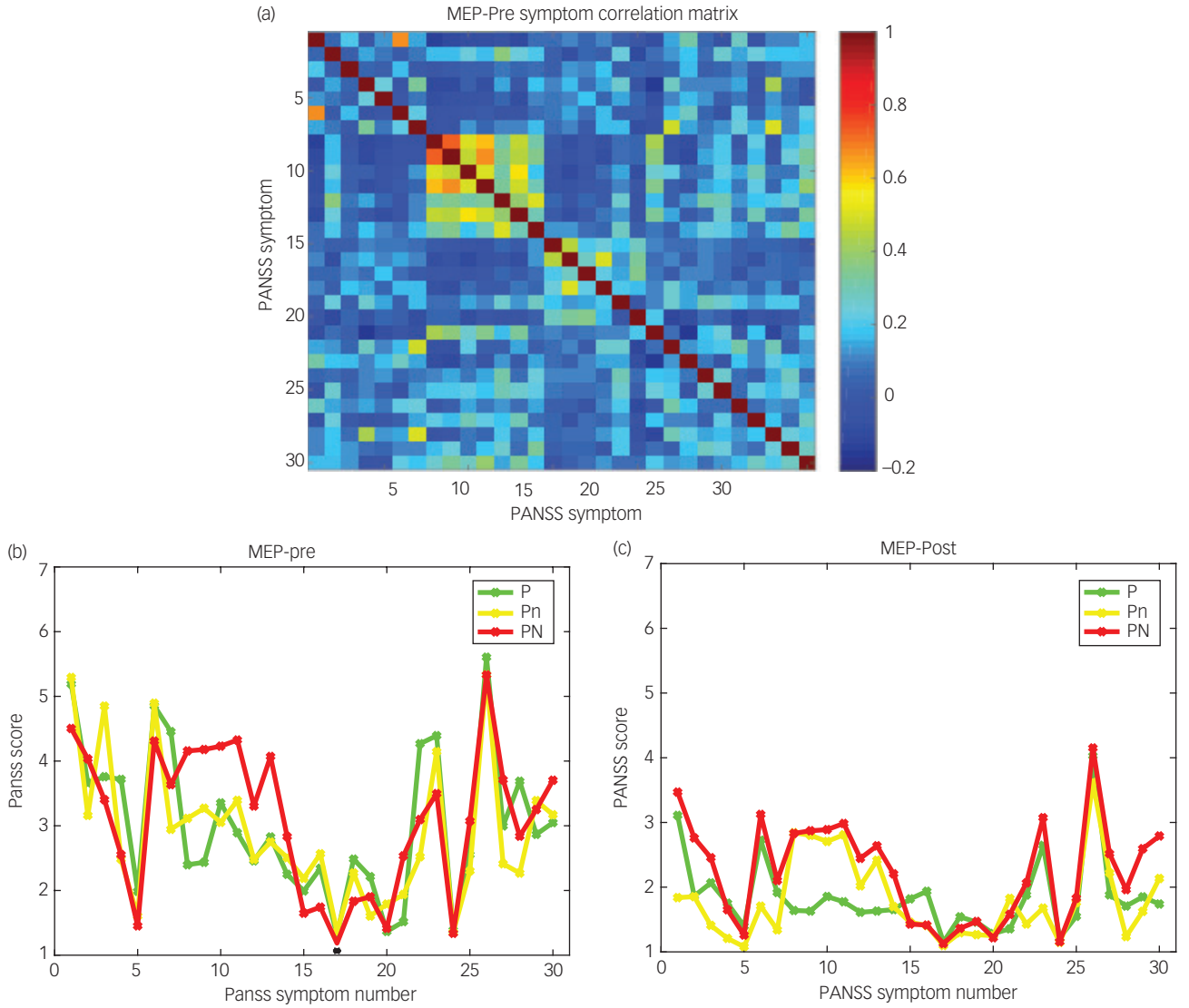


Fig. 1. (a) MEP_Pre symptom correlation matrix. The colour bar indicates the value of the Pearson correlation. In PANSS (14), P1–P7 are Symptoms 1–7, N1–N7 are Symptoms 8–14 and G1–G16 are Symptoms 15–30. The values for each score are 1–7. The symptoms are as follows: Delusions (P1), Conceptual disorganisation (P2), Hallucinations (P3), Hyperactivity (P4), Grandiosity (P5), Suspiciousness/persecution (P6), Hostility (P7), Blunted affect (N1), Emotional withdrawal (N2), Poor rapport (N3), Passive/apathetic social withdrawal (N4), Difficulty in abstract thinking (N5), Lack of spontaneity and flow of conversation (N6), Stereotyped thinking (N7), Somatic concern (G1), Anxiety (G2), Guilt feelings (G3), Tension (G4), Mannerisms and posturing (G5), Depression (G6), Motor retardation (G7), Uncooperativeness (G8), Unusual thought content (G9), Disorientation (G10), Poor attention (G11), Lack of judgment and insight (G12), Disturbance of volition (G13), Poor impulse control (G14), Preoccupation (G15) and Active social avoidance (G16). (b) MEP_Pre average symptom values in the three patient clusters detected by k-means. (c) MEP_Post average symptom values in the three patient clusters detected by k-means.

the k-means analysis. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms (N1–N7=symptoms 8–14). The group labelled Pn (positive and intermediate negative symptoms) had intermediate values for the negative symptoms. The group labelled P (positive and low negative symptoms) had low values for the negative symptoms. Little else differed between these clusters, except that the PN cluster has a smaller value for P1 (delusions) and P3 (hallucinatory behaviour) (Symptoms 1 and 3) and has higher scores for Symptoms 27–30 (G13–G16). The three clusters contained 228, 233 and 226 patients, respectively. The symptom profiles of the groups detected by the k-means algorithm (Fig. 3b) are very similar to the profiles of the corresponding groups detected by the community detection algorithm (Figs. S9 and S10).

Figure 3a shows that there are moderately high correlations between the seven negative symptoms, N1–N7 (Symptoms 8–14). To investigate how the negative symptoms were distributed across the population of FEP patients, the population × symptom data matrix was sorted by the mean value of Symptoms N1–N7. The result shown in Fig. 4a indicates that there is a graded distribution of the negative symptoms across the population of 687 FEP_Pre patients, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible).

Figure 4b and c shows that the distribution of the mean of the average of the negative symptom scores in the FEP_Pre group has a unimodal distribution, in one dimension, rather than a bimodal or multimodal distribution. Thus the most important way in which

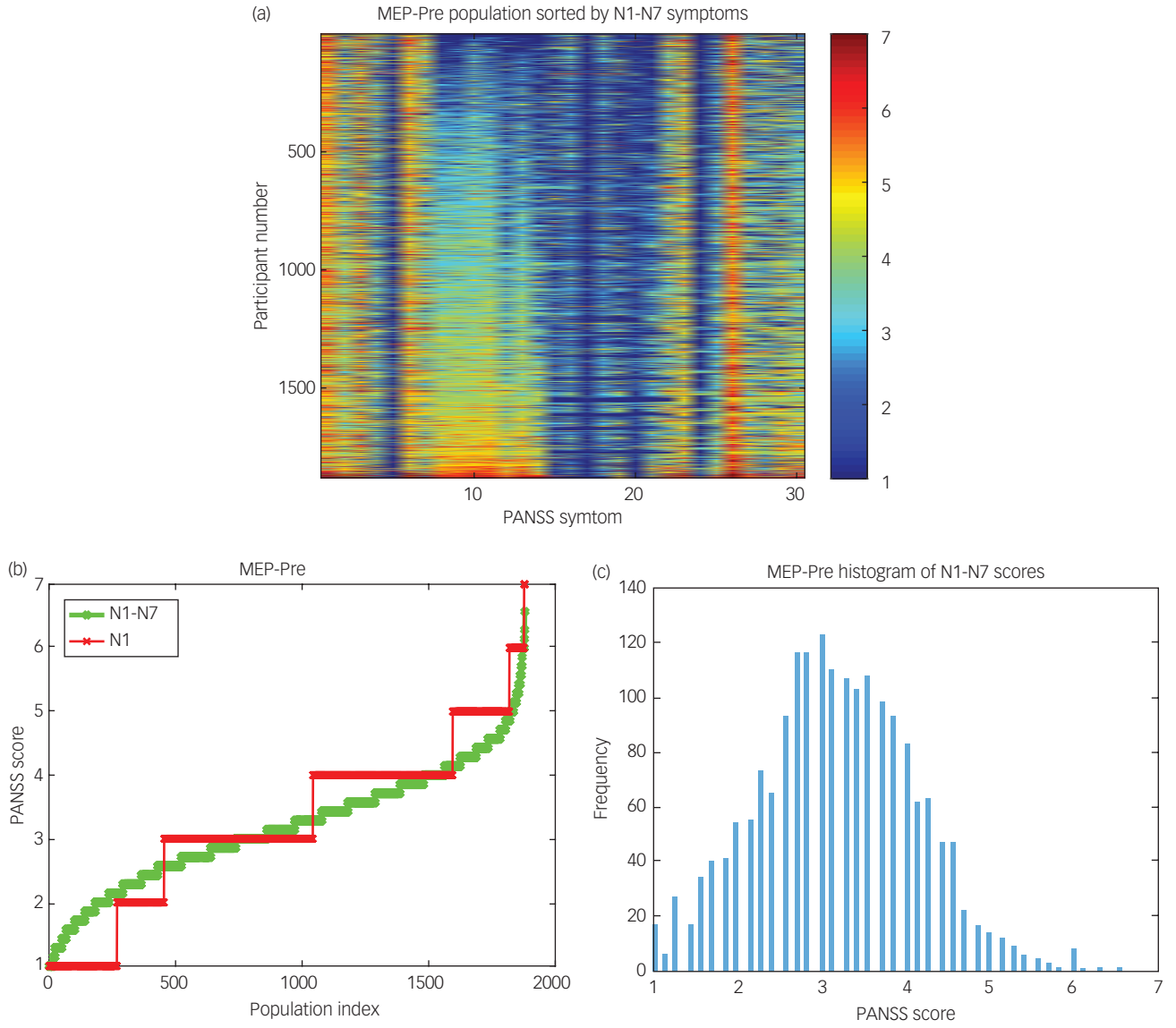


Fig. 2. (a) MEP_Pre group sorted by the average value of the negative symptoms, N1–N7 (Symptoms 8–14). The colour bar on the right shows the PANSS score in the range 1–7. (b) The average score for Symptoms N1–N7 shown sorted by its value in each member of the MEP_Pre population (green). The red plot shows the value for N1. (c) A histogram of the average score for Symptoms N1–N7 in the MEP_Pre group.

the three sub-populations of FEP_Pre patients differed was in the negative symptom scores, which are continuously distributed from 1 to 7.

First-episode group post-treatment (FEP_Post)

The FEP_Post symptom correlation matrix showed, similarly to the FEP_Pre condition illustrated in Fig. 3a, that the negative symptoms (8–14) form a clear cluster (Fig. S12). Three sub-population clusters were again found by the community detection algorithm (Fig. S13), and the three clusters in the FEP_Post dataset that were detected with the k-means algorithm had 288 patients in the PN cluster, 221 in the Pn cluster and 178 in the P cluster.

Figure 3c shows the average symptom values in the three patient clusters for the FEP_Post group. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms (N1–N7=symptoms 8–14). The group labelled Pn (yellow) had high scores for P1–P7 and intermediate values for the negative symptoms. The group labelled P had low values for the negative and positive symptoms. It is notable that the scores on all or most of the symptoms have been decreased by the medication (Fig. 4e). Just as in the FEP_Pre condition, a continuously graded distribution of the negative symptoms across the population of 687 FEP_Post patients was found, with patients at one end of the distribution having scores close to 1 (the minimum score possible) and at the other end close to 7 (the maximum score possible).

Effects of treatment on the PANSS scores

The effects of treatment with antipsychotic drugs on the PANSS symptom scores for the MEP and FEP groups are shown in Fig. 4d and e. Interestingly, the treatment reduced almost all the symptoms (with smaller effects for those close to baseline). The effects on the symptoms were highly similar in the MEP and FEP groups. Moreover, the reductions in the scores of the negative symptoms for the different sub-populations PN, Pn and P produced by the medication were similar (making allowance for the fact that the pre-treatment value was different in the three

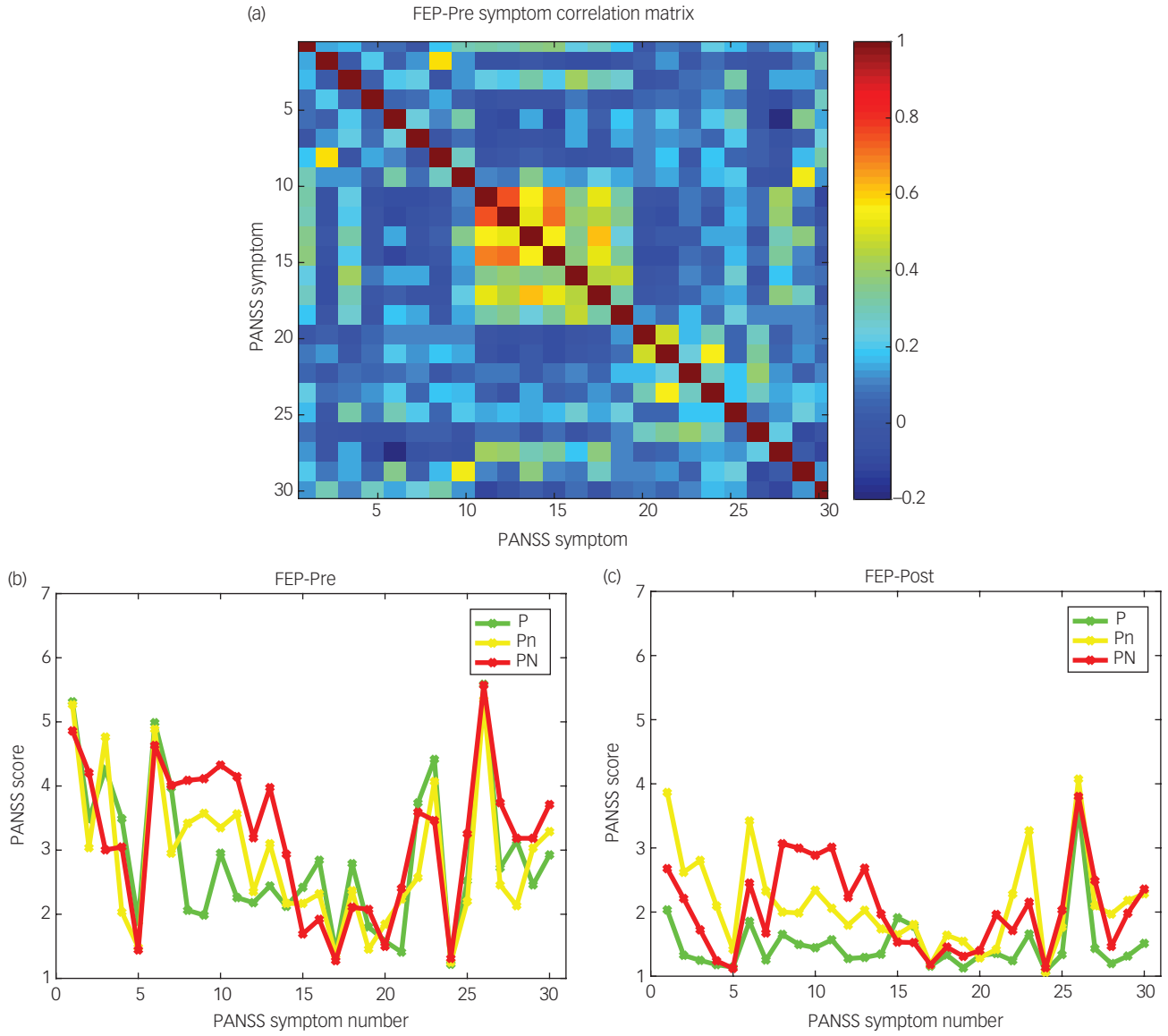


Fig. 3. (a) FEP_Pre symptom correlation matrix. The colour bar indicates the value of the Pearson correlation. In PANSS (14), P1–P7 are Symptoms 1–7, N1–N7 are Symptoms 8–14 and G1–G16 are Symptoms 15–30. The values for each score are 1–7. (b) FEP_Pre average symptom values in the three patient clusters detected by k-means. (c) FEP_Post average symptom values in the three patient clusters detected by k-means.

sub-populations). The treatment consisted of one out of seven randomly assigned drugs (risperidone, quetiapine, perphenazine, olanzapine, haloperidol, aripiprazole and ziprasidone) for a 6-week treatment period. Further details on the effects of the treatments are available from the authors on request and will be provided elsewhere.[16]

Comparison of MEP and FEP groups

It is evident from Fig. 4d and e that in the pre-treatment condition, the FEP and MEP groups had almost identical scores, and this is confirmed by the explicit comparison shown in Fig. S19. It is evident from Fig. 4d and e that in the post-treatment condition, the FEP and MEP groups had almost identical scores (though lower than in the pre-treatment groups), and this is confirmed by the explicit comparison shown in Fig. S20.

Community detection results

The community detection algorithm fast_mo[15] and its modified form fast_mo_sgn[16] applied to the population correlation matrix produced very similar results as shown in the Data supplement (Figs. S1–S14) to those described in the main text using k-means, with the main difference that the numbers of patients in the different clusters or communities differed somewhat.

Factor analysis and multidimensional scaling on the symptom correlation matrices

To provide further insight into how the different PANSS symptoms are related to each other and how they separate from each other, factor analyses and multidimensional scaling were performed on the symptom correlation matrices, as described in the Data

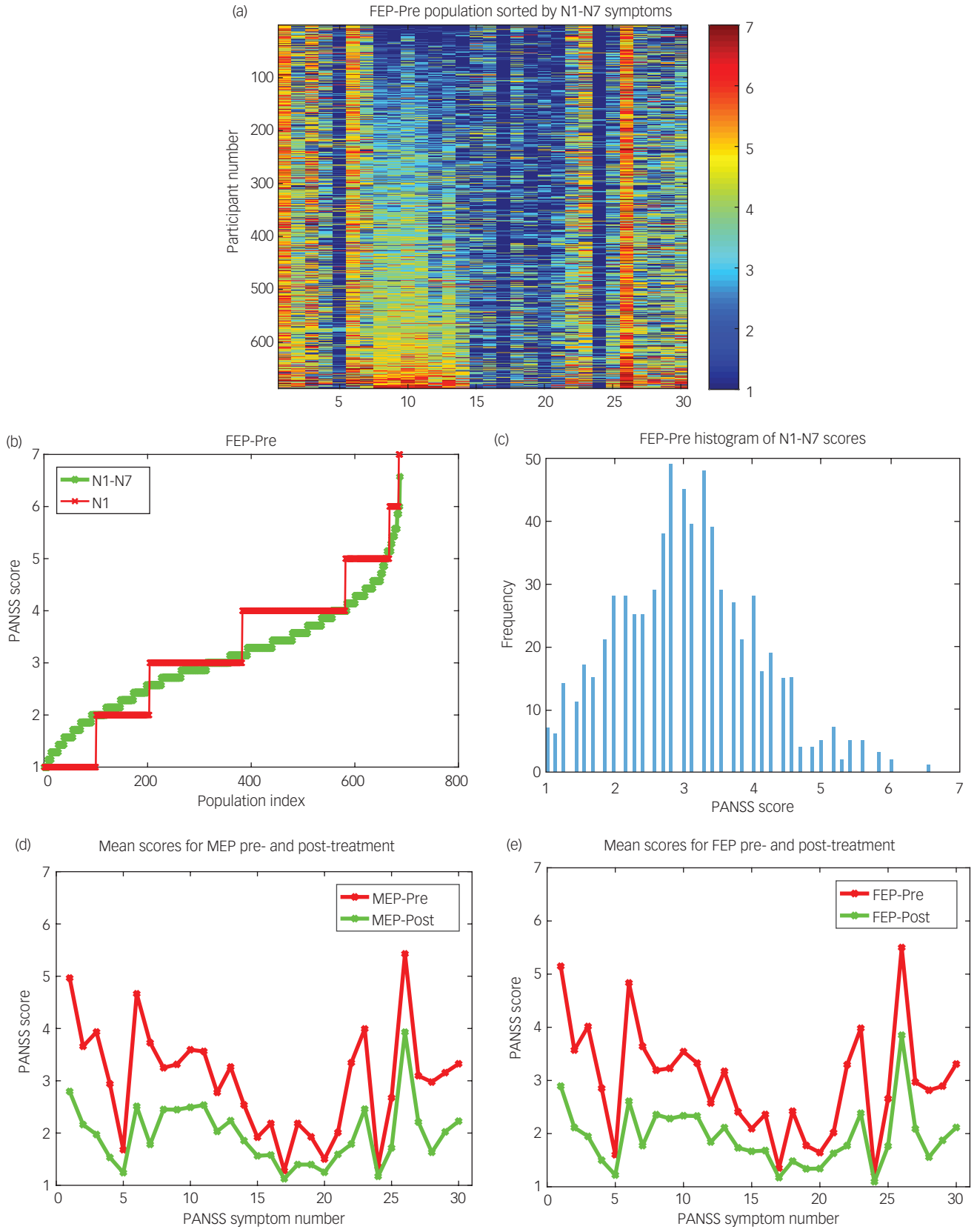


Fig. 4. (a) FEP_Pre group sorted by the average value of the negative symptoms, N1–N7 (Symptoms 8–14). The colour bar on the right shows the PANSS score in the range 1–7. (b) The average score for Symptoms N1–N7 shown sorted by its value in each member of the FEP_Pre population (green). The red plot shows the value for N1. (c) A histogram of the average score for Symptoms N1–N7 in the FEP_Pre group. (d and e) PANSS symptom scores before and after treatment for (d) the MEP and (e) the FEP groups.

supplement. These confirm that the major source of variation between individuals with schizophrenia is in the negative symptoms.

Discussion

The results of the analysis on the datasets with 687 FEP with schizophrenia before treatment (FEP_Pre) and after treatment (FEP_Post), and 1880 MEP before treatment (MEP_Pre) and after treatment (MEP_Post), provide evidence that a major difference between patients is the extent of the negative symptoms (Figs 1–4). There is little evidence for very discrete populations, with instead a gradation in the amount of the negative symptoms found across the patient populations (Figs 1–4). Indeed, the distributions for the mean of the negative symptoms N1–N7 are unimodal (Figs 2 and 4), so there is little support for discrete subtypes of schizophrenia, on the basis of PANSS scores. Indeed, when the k-means and the community detection algorithms did separate the patient populations into three groups, these groups were distinguished mainly by the amount of the negative symptoms, as illustrated in Figs 1, 3, S3, S4, S10 and S11. Figures 4d, e and S19 show that the PANSS symptoms are almost identical before treatment in the FEP and MEP groups. Figures 4d, e and S20 show that the PANSS symptoms are almost identical after treatment in the FEP and MEP groups. The results shown in Fig. 4d and e show that medication decreases the scores on essentially all PANSS symptom scores.

Our finding of a continuous unimodal distribution of negative symptoms is consistent with many cross-sectional studies that have described the heterogeneity of schizophrenia in terms of multiple dimensions (reviewed by Peralta & Cuesta[10]). Our findings indicate that the negative dimension plays a particularly important role in differentiating patients, but do not provide evidence for discrete subtypes of illness. However, our cross-sectional analyses do not address the question of whether or not taking account of persistence of symptoms or treatment response might help identify discrete subtypes. Some previous studies that have taken account of the tendency of symptoms to persist have identified a subtype of illness designated deficit schizophrenia, characterised by persistent primary negative symptoms.[17,18] Other studies, reviewed by Gillespie *et al*,[19] have concluded that treatment-resistant schizophrenia might be categorically distinct from treatment-responsive schizophrenia.

Evidence from many treatment trials, reviewed by Barry *et al*,[20] indicates that negative symptoms tend to be poorly responsive to treatment. However, we found in our very large patient sample that positive and negative symptoms reduced to a similar degree during treatment. An issue that must be considered is the characteristics of the patient sample, including selection criteria. It is illustrative to compare our findings with Tollefson *et al*[21] in a trial of olanzapine compared with haloperidol in a large sample of approximately 2000 patients. (Although the focus of Tollefson *et al* study was on the differences between responses to the two different antipsychotics, those differences were small and for the purpose of this discussion the differences between the antipsychotics is of limited relevance.) Tollefson *et al* reported a reduction of six points in PANSS negative symptom scores after 6 weeks of treatment with olanzapine and five points after 6-weeks treatment with haloperidol. We found a reduction in PANSS negative symptom scores of 8.2 points in MEP and 8.3 points in FEP cases. It is possible that the fact that we excluded cases with treatment resistant illness, whereas Tollefson *et al* did not, contributed to the difference in the magnitude of the reduction produced by antipsychotics. In addition, it is possible that social, cultural or ethnic influences contributed to the greater reduction in negative symptoms in our sample. We discuss this in greater detail in the Data supplement in the 'Social, cultural and ethnic influences on response to treatment' section.

It should be noted that we did not attempt to distinguish primary from secondary negative symptoms. Our finding of a significant correlation between the reduction in positive symptoms and the reduction in negative symptoms raises the possibility that a proportion of the negative symptoms reported in our study might be secondary negative symptoms which are expected to decrease during treatment. However, the results of an analysis of covariance presented in the Data supplement show that the reduction in the negative symptoms produced by the medication was in part independent of the reduction in the positive symptoms.

Furthermore, it is noteworthy that we found MEP and FEP cases exhibited similar decreases in symptom severity during 6-weeks treatment. In contrast, other moderately large studies have reported that FEP cases show a greater reduction in symptoms during treatment than MEP cases.[22] In a longitudinal study extending over the first 5 years of illness, Lieberman *et al*[23] found that treatment response diminished with increasing number of relapses. It is possible that the exclusion of participants with treatment-resistant illness from our MEP sample contributed to our finding of a relatively large reduction of negative symptoms in MEP cases.

Further analyses showed the following

The negative symptoms (N1–N7) are moderately well correlated with each other (Figs 1a and 3a). As demonstrated in the Data supplement, the results of applying the community detection algorithm fast_mo_sgn to the symptoms showed that in the MEP and FEP groups, both Pre- and Post-treatment, four symptom communities were detected (with one minor exception). One community generally included most of the positive symptoms and another community included most of the negative symptoms (Figs S1, S5, S8 and S12). The analysis of symptom communities identified a community of symptoms dominated by anxiety and depression (G6) in both FEP (Fig. S8) and MEP (Fig. S1) groups, both before and after treatment. This is consistent with many studies that have reported a depression syndrome in schizophrenia (see the review by Peralta & Cuesta[10]). Furthermore, in the MEP_Pre cases, there was a community of symptoms reflecting disorganisation of thought and behaviour, together with cognitive symptoms such as poor attention, resembling the disorganisation syndrome identified in symptomatic cases of chronic schizophrenia by Bilder *et al*[9] and Liddle.[7] Further analysis of just the negative symptoms found that they could be separated into three communities (N1, N2 and N4; N3 and N6; and N5 and N7), and that these symptom communities were all similarly reduced by medication (Data supplement).

The results shown in Figs S19 and S20 show that the PANSS symptoms are very similar between FEP and MEP groups without medication; and that the medication produces changes in both groups which make the FEP_Post and MEP_Post groups very similar after medication. These findings have major implications for 'staging' hypotheses of schizophrenia. Any 'staging' hypothesis would need to carefully factor out the effects of treatment for its results to be interpretable. It may be that very long-term patients with schizophrenia do show symptom differences.[5] But if so, that would be related to the long-term nature of the disease not clearly evident in the MEP and FEP groups described here.

K-means cluster analysis can be used to divide the patients into 2–4 sub-populations (with results for three sub-populations presented), which differ from each other mainly in the extent of the negative symptoms. Community detection analysis (using fast_mo_sgn)[16] (modified from[15]) provides a very similar set of three sub-populations of both the MEP and FEP group, though the numbers of patients in each population are not identical. To the extent that the distribution of the negative symptoms is in part continuous, the cluster/community detection may not imply that the sub-populations are very

discrete. Instead, a continuous distribution of the negative symptom scores from 1 to 7 across the population, as illustrated in Figs 1–4, needs to be considered. One way to assess patients clinically might be on the mean of the seven negative symptom PANSS scores, and clinicians may wish to consider different treatments with the graded distribution across the population in mind, with for example some findings becoming available from Lu *et al* (in preparation).[16]

The finding that the main difference in the PANSS scores of different patients is in the extent of the negative symptoms is of considerable interest. If a k-means analysis is performed using only the values of the negative symptoms, then three population communities are found which differ in the mean value of the negative symptoms (Figs S4 and S11). This adds to the evidence that the values of the negative symptom scores are an important factor in producing different sub-populations or communities, and moreover that what distinguishes these communities includes a difference in the mean negative symptom scores. Furthermore, the evidence in the Data supplement indicates that the negative symptoms are independent of the positive and general symptoms (i.e. the correlations between these symptoms pre-treatment is close to zero), so that there must be at least two partly independent underlying brain processes that are involved.

Subtypes of schizophrenia are no longer included in DSM-5, and this large study with unmedicated patients is of interest in that context, for no highly discrete subtypes of patients were detected in this study, but instead there was a continuous distribution of the mean value of the negative symptoms in both the FEP and MEP populations. The subtypes of symptoms illustrated in Figs S1, S5, S8 and S12 do show how a modern community detection algorithm clusters together different symptoms, with one symptom community including, for example, most of the positive symptoms, and a second community including most of the negative symptoms.

The present investigation adds to previous analyses of subtypes of schizophrenia (e.g.[11,24–26]) by using a very much larger sample of patients (2567) than in any previous investigation; by investigating the symptoms in the same patients in both the unmedicated and medicated states; by comparing the symptoms in FEP and MEP and by including a more powerful machine learning, community detection, approach. This investigation differs from the analysis of a partly overlapping FEP patient population by Lu *et al* (in preparation)[16] in that we used here the same participants unmedicated and medicated to understand the effects of the medication, which may produce clearer results than comparing two different groups when unmedicated or not; in the use of more classical clustering approaches to define how the symptoms cluster and how the members of the population cluster; and in the emphasis on the continuous unimodal distribution of the negative symptoms which differs between individuals, instead of focusing on communities that are thought to be distinct because they have been categorised separately by a community detection algorithm which looks for distinct communities.[15]

Apart from the large sample size in this study, a strength is that the patients came from several centres, and that similar results were obtained from the different centers.[16] Another strength is assessment of the PANSS when unmedicated and after 6 weeks of medication in the same patients. Several limitations may be mentioned. PANSS data were used to identify the communities in the present investigation and the communities may become further refined when further measures of differences between patients are included. Some PANSS items, such as difficulty in abstract thinking (N5) and poor attention (G11), include what may be different aspects of impoverished and disorganised thinking within single items, and separate assessment might lead to a richer categorisation.[7,27] Further we suggest that the classification of possible subtypes of schizophrenia could be enhanced by also including information from functional neuroimaging and from genetics.

In conclusion, in a very large sample of cases of schizophrenia, we have demonstrated that variation in the negative symptoms is an important aspect of the variation between patients with schizophrenia. The severity of the negative symptoms has a continuous unimodal distribution supporting a dimensional description of the heterogeneity of the illness, at least with respect to the negative symptoms. Furthermore, in this sample from which treatment resistant cases were excluded, the severity of the negative symptoms decreased substantially during 6 weeks of antipsychotic treatment in both FEP and MEP cases.

Edmund T. Rolls, MA, DPhil, DSc, Department of Computer Science, University of Warwick, Coventry, UK; Oxford Centre for Computational Neuroscience, Oxford, UK; **Wenlian Lu,** PhD, Centre for Computational Systems Biology, Fudan University, Shanghai, PR China; **Lin Wan**, PhD, National Center for Mathematics and Interdisciplinary Sciences, The Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, PR China; **Hao Yan**, MD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; **Chuanyue Wang**, MD, Beijing Anding Hospital, Capital Medical University, Beijing, PR China; **Fude Yang**, M.M., Beijing HuiLongGuan Hospital, Peking University, Beijing, PR China; **Yunlong Tan**, Beijing HuiLongGuan Hospital, Peking University, Beijing, PR China; **Lingjiang Li**, MD, Institute of Mental Health, The Second Xiangya Hospital of Central South University, Changsha, PR China; **Chinese Schizophrenia Collaboration Group** (see Data supplement); **Hao Yu**, PhD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; **Peter F. Liddle**, MBBS, MRCPsych, PhD, Centre for Translational Neuroimaging, Institute of Mental Health, Division of Psychiatry & Applied Psychology, University of Nottingham, Nottingham, UK; Sir Peter Mansfield MR Centre, University of Nottingham, Nottingham, UK; **Lena Palaniyappan**, MBBS, PhD, Department of Psychiatry, University of Western Ontario, London, Ontario, Canada; Department of Medical Biophysics, University of Western Ontario, London, Ontario, Canada; Robarts & Lawson Health Research Institutes, London, Ontario, Canada; **Dai Zhang**, MD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; Peking-Tsinghua Joint Center for Life Sciences/PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China; **Weihua Yue**, MD, Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, PR China; Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders, Peking University, Beijing, PR China; **Jianfeng Feng**, PhD, Department of Computer Science, University of Warwick, Coventry, UK

Correspondence: Jianfeng Feng, Centre for Computational Systems Biology, School of Mathematical Sciences, Fudan University, Shanghai, China; Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK. E-mail: jianfeng64@gmail.com; Professor Weihua Yue, The Sixth Hospital, Peking University, Beijing, 100191, China. Email: dryue@bjmu.edu.cn; Edmund Rolls, University of Warwick, Coventry, UK. Email: Edmund.Rolls@warwick.ac.uk.

First received 4 May 2017, final revision 31 Jul 2017, accepted 23 Sep 2017

Funding

J.F.F. is a Royal Society Wolfson Research Merit Award holder, partially supported by the NCMIS of the Chinese Academy of Sciences (CAS) and Key Program of National Natural Science Foundation of China (NSFC) grant (No. 91230201). W.H.Y. was supported by grants from the National Key Research and Development Program (2016YFC1307000), National Key Technology R&D Program of China (2015BAI13B01) and National High Technology Research and Development Program of China (2008AA02Z401, 2009AA022702). L.W. is supported by the Strategic Priority Research Program of the CAS (XDB13050000), the NCMIS of the CAS, the LSC of the CAS, the Youth Innovation Promotion Association of the CAS, and NSFC grant (No.11571349). S.X.G. is supported by the NSFC grant (No. 11271121), Program for New Century Excellent Talents in University (NCET-13-0786) grant, and Natural Science Foundation of Hunan Province (2015JJ1010). P.F.L. is supported by MRC grant MR/K020803/1 for an investigation of brain function in schizophrenia. L.P. is supported by the Academic Medical Organization of Southwestern Ontario (AMOSO) and Canadian Institute of Health Research [CIHR Grant No. (377213/ 201610PJT)].

The work of the Chinese Schizophrenia Collaboration Group (contact Prof W Yue, dryue@bjmu.edu.cn) is fully acknowledged. The members and their affiliations are listed in the data supplement.

References

1 Tsuang MT, Lyons MJ, Faraone SV. Heterogeneity of schizophrenia. Conceptual models and analytic strategies. *Br J Psychiatry* 1990; **156**: 17–26.

2 Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–7.

3 Rolls ET. *Cerebral Cortex: Principles of Operation*. Oxford University Press, 2016.

4 Rolls ET, Loh M, Deco G, Winterer G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat Rev Neurosci* 2008; **9**: 696–709.

5 Li T, Wang Q, Zhang J, Rolls ET, Yang W, Palaniyappan L, et al. Brain-wide analysis of functional connectivity in first-episode and chronic stages of schizophrenia. *Schizophr Bull* 2017; **43**: 436–48.

6 Liddle PF. Syndromes of schizophrenia on factor analysis (letter). *Br J Psychiatry* 1992; **161**: 861.

7 Liddle PF. The symptoms of chronic schizophrenia. A re-examination of the positive–negative dichotomy. *Br J Psychiatry* 1987; **151**: 145–51.

8 Owen MJ, Sawa A, Mortensen PB. Schizophrenia. *Lancet* 2016; **388**: 86–97.

9 Bilder RM, Mukherjee S, Rieder RO, Pandurangi AK. Symptomatic and neuropsychological components of defect states. *Schizophr Bull* 1985; **11**: 409–19.

10 Peralta V, Cuesta MJ. A dimensional and categorical architecture for the classification of psychotic disorders. *World Psychiatry* 2007; **6**: 100–1.

11 Strauss JS, Bartko JJ, Carpenter WT Jr. The use of clustering techniques for the classification of psychiatric patients. *Br J Psychiatry* 1973; **122**: 531–40.

12 Newman ME, Girvan M. Finding and evaluating community structure in networks. *Physical Rev E* 2004; **69**: 026113.

13 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). APA, 1984.

14 Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987; **13**: 261–76.

15 Le Martelot E, Hankin C. Fast multi-scale detection of relevant communities in large-scale networks. *Comput J* 2013; **56**: 1136–50.

16 Lu W, Wan L, Rolls ET, Liddle PF, Ma L, Yan H, et al. Novel subtyping of schizophrenia predicts response to antipsychotics. 2018. [paper in preparation]

17 Kirkpatrick B, Galderisi S. Deficit schizophrenia: an update. *World Psychiatry* 2008; **7**: 143–7.

18 Ahmed AO, Strauss GP, Buchanan RW, Kirkpatrick B, Carpenter WT. Are negative symptoms dimensional or categorical? Detection and validation of deficit schizophrenia with taxometric and latent variable mixture models. *Schizophr Bull* 2015; **41**: 879–91.

19 Gillespie AL, Samanaite R, Mill J, Egerton A, MacCabe JH. Is treatment-resistant schizophrenia categorically distinct from treatment-responsive schizophrenia? A systematic review. *BMC Psychiatry* 2017; **17**: 12.

20 Barry SJ, Gaughan TM, Hunter R. Schizophrenia. *BMJ Clin Evid* Jun 2012 (http://clinicalevidence.bmj.com/x/systematic-review/1007/overview.html). Accessed 26 April 2017.

21 Tollefson GD, Beasley CM Jr, Tran PV, Street JS, Krueger JA, Tamura RN, et al. Olanzapine versus haloperidol in the treatment of schizophrenia and schizoaffective and schizophreniform disorders: results of an international collaborative trial. *Am J Psychiatry* 1997; **154**: 457–65.

22 Jager M, Riedel M, Messer T, Laux G, Pfeiffer H, Naber D, et al. Psychopathological characteristics and treatment response of first episode compared with multiple episode schizophrenic disorders. *Eur Arch Psychiatry Clin Neurosci* 2007; **257**: 47–53.

23 Lieberman JA, Koreen AR, Chakos M, Sheitman B, Woerner M, Alvir JM, et al. Factors influencing treatment response and outcome of first-episode schizophrenia: implications for understanding the pathophysiology of schizophrenia. *J Clin Psychiatry* 1996; **57** (Suppl 9): 5–9.

24 Carpenter WT Jr, Stephens JH. An attempted integration of information relevant to schizophrenic subtypes. *Schizophr Bull* 1979; **5**: 490–506.

25 Carpenter WT Jr, Bartko JJ, Carpenter CL, Strauss JS. Another view of schizophrenia subtypes. A report from the international pilot study of schizophrenia. *Arch Gen Psychiatry* 1976; **33**: 508–16.

26 Strauss JS, Carpenter WT Jr, Bartko JJ. The diagnosis and understanding of schizophrenia. Part III. Speculations on the processes that underlie schizophrenic symptoms and signs. *Schizophr Bull* 1974; **11**: 61–9.

27 Baxter RD, Liddle PF. Neuropsychological deficits associated with schizophrenic syndromes. *Schizophr Res* 1998; **30**: 239–49.

BJPO/2017/005058, doi: 10.1192/bjpo.bp.117.005058

Data supplement

Individual differences in schizophrenia

Edmund T Rolls[1,2], Wenlian Lu[3], Lin Wan[4], Hao Yan[5,6], Chuanyue Wang[7], Fude Yang[8], Yun-Long Tan[8], Lingjiang Li[9], Chinese Schizophrenia Collaboration Group[10], Hao Yu[5,6], Peter F Liddle[11,12], Lena Palaniyappan[13], Dai Zhang[5,6,14], Weihua Yue[5,6*], Jianfeng Feng[1,3*]

[1] Department of Computer Science, University of Warwick, Coventry, UK;

[2] Oxford Centre for Computational Neuroscience, Oxford, UK

[3] Centre for Computational Systems Biology, Fudan University, Shanghai, PR China;

[4] National Center for Mathematics and Interdisciplinary Sciences, and the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, PR China;

[3] College of Mathematics and Computer Science, Key Laboratory of High Performance Computing and Stochastic Information Processing (Ministry of Education of China), Hunan Normal University, Changsha, PR China;

[5] Institute of Mental Health, the Sixth Hospital, Peking University, Beijing, 100191, China;

[6] Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders (Peking University), Beijing, 100191, China;

[7] Beijing Anding Hospital, Capital Medical University, Beijing 100096, PR China.

[8] Beijing HuiLongGuan Hospital, Peking University, Beijing 100096, China;

[9] Institute of Mental Health, The Second Xiangya Hospital of Central South University, Changsha 410011, China;

[10] Chinese Schizophrenia Collaboration Group: see Supplementary Material 2.

[11] Centre for Translational Neuroimaging, Institute of Mental Health, Division of Psychiatry & Applied Psychology, University of Nottingham, Nottingham, UK;

[12] Sir Peter Mansfield MR Centre, University of Nottingham, Nottingham, UK;

[13] Departments of Psychiatry & Medical Biophysics, University of Western Ontario, London, Ontario; and Robarts & Lawson Health Research Institutes, London, Ontario, Canada.

[14] Peking-Tsinghua Joint Center for Life Sciences/PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China.

Overview

This Supplementary material presents further details of the results in the main paper, in some cases using a modern community detection method for separating patients into different "subtypes", subpopulations, clusters, or communities. The community detection method used is that described by Le Martelot and Hankin (1), using the algorithm fast_mo.m. This is described below, as is a small modification intended to help the community detection deal better with negative values in the correlation matrix, though with the current dataset it was found that this modification made little or no difference. The modified code, fast_mo.sgn.m (2) is available on request to the corresponding author. The community detection method used here is essentially that of Le Martelot and Hankin (1).

Community detection algorithm for graphs with negative weights

The community detection algorithm is based on the idea of modularity as described by Newman and Girvan (3) and a modification by Traag and Bruggeman (4) which takes the negative edges into consideration. In the present context, details of which are provided elsewhere (2), a negative edge might be a negative correlation between a pair of symptoms in a symptoms cross-correlation matrix, or between a pair of individuals in the population cross-correlation matrix. Consider a weighted graph $\mathrm{G} = [\mathrm{V}, \mathrm{E}, \mathrm{W}]$ with V the node set, E the edge set and $\mathrm{W} = [w_{ij}]_{i,j \in V}$ the weight matrix. In detail, an edge from node i to node j exists, namely, $\mathrm{e(i,j)} \in \mathrm{E}$ if and only if $w_{ij} \neq 0$. Here, we consider a symmetrical graph, i.e., W is a symmetrical matrix. We emphasize that this method can be extended to directed graphs with asymmetrical weights easily. Consequently, the graph G can be divided into two graphs: G^+ and G^-, which comprise the same node set V, the positive/negative edge subsets, and the positive/negative weight matrices respectively. Both positive and negative edges are considered. Thus, the edge set E is divided into two subsets: the positive edge subset $E^+ = \{\mathrm{e(i,j)} \in \mathrm{E}: w_{ij} > 0\}$ and the negative edge subset $E^- = \{\mathrm{e(i,j)} \in \mathrm{E}: w_{ij} < 0\}$. Thus, define the positive weighted degrees and negative weighted degrees of each node as follows respectively

$$k_i^+ = \sum_{e(i,j) \in E^+} w_{ij}\,, k_i^- = \sum_{e(i,j) \in E^-} w_{ij}\,, \mathrm{i} \in \mathrm{V}$$

Essentially, nodes in the same community have more positive edges and fewer negative edges than nodes between communities. The extended modularity is the sum of the differences between the fraction of edge weights that fall within communities, minus the expected value of the same quantity if edges fall as random graphs, considering the positive and negative edges, without regard for the community structure, as given in Equation (1) (4):

$$Q_M(W, \delta) = \frac{1}{2m} \sum_{i,j \in V} [w_{ij} - (p_{ij}^+ - p_{ij}^-)] \cdot \delta(i,j). \qquad (1)$$

Here m is the total of the absolute values of the weighted matrix W, the community structure function $\delta(\mathrm{i,j})$returns one if nodes i and j belong to the same community, and zero otherwise (5), and $p_{ij}^{\pm}$ stands for

the coupling probability from node i to node j in random graphs that follows the same weight distribution in $G^{\pm}$ resepectively. In detail

$$p_{ij}^{+} = \frac{k_i^{+} k_j^{+}}{m^{+}}, p_{ij}^{-} = \frac{k_i^{-} k_j^{-}}{m^{-}}$$

where $m^{\pm}$ are the sum of the absolute values of the weighted matrices of the positive and negative graph $G^{\pm}$ respectively. The higher the value of Q_M, the better is the community division.

Given the number of communities, denoted by κ, the fast community detection algorithm that we used (1) starts with a state in which each vertex is randomly organized into κ communities (1). (The Matlab code 'fast_mo.m' that we used in this study was downloaded from http://www.elemartelot.org/index.php/programming/cd-code, and was modified to work as described here with negative edges.) We repeatedly change the community label of each vertex (in a random order) and choose at each step the join that results in the greatest increase in Q_M until no improvement occurs. This process is repeated 100 times and the best result (with the largest Q_M) is taken (1). We take different values of k and repeat this algorithm and detect the community structure with the largest Q_M over different numbers of communities. It is noted that the community method can subtract the background noise in the Q_M function, and this makes it potentially more robust and accurate than classical clustering methods, for instance, k-means. In comparison to the original fast_mo algorithm (1) in which the number of communities is calculated in an unsupervised way by a hierarchical clustering method, both community structures that have intersections but local maximum values of Q_M can be probed by our algorithm (fast_mo_sgn) (2) but may be missed by that fast_mo (1).

The disagreement between two community structures, $\delta_1(\cdot,\cdot)$ and $\delta_2(\cdot,\cdot)$, is measured as follows (2):

$$\mathrm{dis}(\delta_1, \delta_2) = \frac{1}{N(N-1)/2} \sum_{i>j} C_{\delta_1,\delta_2}(i,j)$$

with

$$C_{\delta_1,\delta_2}(i,j) = \begin{cases} 1 & \delta_1(i,j) = \delta_2(i,j) \\ 0 & \delta_1(i,j) \neq \delta_2(i,j) \end{cases}$$

where N stands for the number of the nodes in the graph.

Participants – further details

Two samples comprising a total of 2,567 patients with schizophrenia were diagnosed according to DSM-IV criteria (6) and the symptoms were assessed using the Positive and Negative Syndrome Scale for Schizophrenia (7). The first study sample included 687 first-episode, drug-naïve patients (FEP) with schizophrenia. The second study sample included 1880 multi-episode patients (MEP) with schizophrenia

recruited from inpatients at the same centers. The patients were recruited from inpatients and outpatients from the mental health departments of four institutes: the Sixth Hospital of Peking University, the Second Xiangya Hospital of Central South University, Beijing Anding Hospital, and Beijing HuiLongGuan Hospital with institutional approval of the investigation and informed consent of the patients.

Patients entered the study according to the following criteria: 1) 18-45 years old; 2) Han Chinese descendants; 3) a diagnosis of schizophrenia, using the Structured Clinical Interview of the DSM-IV (SCID); 4) physically healthy and had all laboratory parameters within normal limits; 5) total score of the PANSS at baseline was more than 60, and at least 3 positive items were scored more than 4. The consensus diagnoses were made by at least two experienced psychiatrists on the basis of structured interviews (SCID) with patients and families and review of medical records. The training of the psychiatrists for this investigation included training in research protocols, standard diagnostic criteria, and other scales for assessment of symptoms and side effects, as well as videos for standardized scale assessment. The specific training took ten days. The PANSS was administered by the psychiatrists. The MEP patients had several previous episodes of schizophrenia, with the average duration of the illness 7.97 years.

Patients were excluded using the following criteria: 1) a diagnosis of schizoaffective disorder, mental retardation, or other cognitive disorders; 2) a history of serious adverse reactions to the proposed treatments; 3) a history of treatment resistance, defined by the persistence of severe symptoms despite adequate trials of one of the proposed treatments or prior treatment with clozapine; 4) pregnant or breast-feeding; 5) a serious and unstable medical condition. The dropout rate was 11.99%.

RESULTS with the community detection analysis

MEP group

We start with the Multi-Episode (MEP) group, as this is the larger group, with 1880 patients with multiple episode schizophrenia, allowing very robust statistical analysis. We start with the pretreatment results when no medication was present (MEP_Pre).

MEP_Pre

Fig. 1a of the main paper shows the MEP_Pre symptom correlation matrix. The community detection method detected four symptom communities, as illustrated in Fig. S1. Symptom community 4 included the majority of the positive symptoms (P1 P4 P5 P6 P7) and G8 G9 G12 G14. Symptom community 1 included the majority of the negative symptoms (N1 N2 N3 N4 N6) and G7 G13 G16. Symptom community 2 included

conceptual disorganization (P2), difficulty with abstract thinking (N5), stereotyped thinking (N7) together with several general symptoms (G5 G10 G11 G15) that largely reflect disorganized behavior. Symptom community 3 comprised P3 (hallucinatory behaviour) together with 5 affective symptoms, G1 G2 G3 G4 G6. Note that the order of the symptoms along the axes has been reordered to reflect the communities detected.

Fig. S2 shows the three patient subpopulation clusters that were detected with the community detection algorithm, with 770, 318, and 792 patients in each cluster. It is evident that the negative symptoms (8-14) differ between the 3 patient groups, with population 1 having the highest scores for the Negative symptoms, population 2 intermediate scores, and population 3 the lowest scores.

Fig. S3 shows the average symptom values in the three patient clusters for the MEP_Pre group using the community detection algorithm. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms N1-N7 = symptoms 8-14). The group labelled Pn (positive and intermediate negative symptoms) had intermediate values for the negative symptoms. The group labelled P (positive and low negative symptoms) had low values for the negative symptoms, especially N1 and N2. Little else differed between these clusters, except that the PN cluster has a higher value for P2 (symptom 2, conceptual disorganization), and smaller value for P6 (symptom 6, suspiciousness/persecution); and the P cluster had relatively high scores on P1, P3-P7, and symptoms 22 and 23 namely G8 (uncooperativeness) and G9 (unusual thought content). A comparison of Fig. 1b and Fig. S3 shows that the community detection method (Fig. S3) produced almost the same mean values for the symptoms of the three groups as the k-means analysis (Fig. 1b), although the numbers of patients in each group were somewhat different. Further, the separation into different populations reflected the mean values of the negative symptoms, which were for community P 2.63, for Pn 2.93, and for PN 3.86.

Fig. S4 shows the community detection results based only on the 7 negative symptoms. Again, three populations were detected, which differed in having values for the negative symptoms that were high (PN), intermediate (Pn), or low (P). The fact that the community detection method found three populations when only the negative symptoms were considered (the same number as when all the symptoms were considered, see Figs. S2 and S3) strengthens the evidence that an important factor in determining how the populations san be separated lies in the values of the negative symptoms.

MEP_Post

We now consider the multiepisode patients in the post-treatment, medicated, state (MEP_Post). The rationale for the different analyses has been described for the MEP_Pre group, and the remainder of the analyses are presented succinctly.

Fig. S5 shows the MEP_Post symptom correlation matrix rearranged according to the 4 communities detected. Community 4 included all the positive symptoms, and community 1 the majority of the negative symptoms. In this medicated states, the positive symptoms form a more uniform cluster than in the unmedicated state where positive symptoms were included in several different communities. The same is the case for the negative symptoms in the medicated state.

Fig. S6 shows the three subpopulation clusters in the MEP_post dataset that were detected with the community detection algorithm, with the number of patients in the three clusters as follows; PN=921, Pn = 19, and P= 940. Of course these post-treatment subpopulation analyses did not necessarily place the same patients in the three groups that were detected pre-treatment, and indeed the interpretation of these analyses is that even post-treatment, the main factor that enables patients to be separated into 3 subpopulations is the magnitude of the negative symptoms.

Fig. S7 shows the average symptom values in these three patient clusters for the MEP_Post group. The group labelled PN (positive and high negative symptoms) had high values for the negative symptoms N1-N7 = symptoms 8-14). The group labelled Pn (positive and intermediate negative symptoms) had almost as high values for the negative symptoms, and lower values for the positive symptoms. The group labelled P (positive and low negative symptoms) had low values for the negative symptoms. A comparison of Fig. 1c and Fig. S7 shows that the community detection method (Fig. S7) produced almost the same mean values for the symptoms of the three groups as the k-means analysis (Fig. 1c), although the numbers of patients in each group were different, and the Pn group was less different from the PN group with the community detection method (Fig. S7) than the k-means method (Fig. 1c).

FEP group

We now analyse the data from the first episode group of patients with schizophrenia (the FEP group), with 687 patients, using the community detection method. We start with the pretreatment results (FEP_Pre).

FEP_Pre

Fig. S8 shows the FEP_Pre symptom correlation matrix rearranged by the 4 main communities detected with community detection. Community 2 included the majority of the negative symptoms but also P2 (conceptual disorganization), and community 1 the majority of the positive symptoms. Comparison with the MEP_Pre case (Fig. S1) shows that in the FEP-Pre patients, the positive symptoms are more within the same community, as are the negative symptoms.

Fig. S9 shows the three population communities that were detected, with 264, 136, and 287 patients in each cluster. The order of the populations in this Figure is PN, Pn, then P.

Fig. S10 shows the average symptom values in the three patient clusters for the FEP_Pre group. A comparison of Fig. 3a and Fig. S10 shows that the community detection method (Fig. S10) produced almost the same mean values for the symptoms of the three groups as the k-means analysis (Fig. 3a), although the numbers of patients in each group were a little different, and there was a small difference in P3 and P4. Further, the separation into different populations reflected the mean values of the negative symptoms, which were for community PN 3.77, for Pn 3.22, and for P 2.34.

To analyze whether there are communities of patients when just the negative symptom scores are considered (N1-N7), the community detection analysis was repeated using only the scores for N1-N7. The results in Fig. S11 show that three clusters are found: one with all the negative symptoms high (PN); one with

most high apart from N4 and N5 (Pn); and a third community of patients with low values for most of the negative symptoms (P). The fact that the community detection method found three populations when only the negative symptoms were considered (the same number as when all the symptoms were considered, see Figs. S9 and S10) strengthens the evidence that an important factor in determining how the populations can be separated lies in the values of the negative symptoms.

FEP_Post

We now consider the first episode patients in the post-treatment, medicated, state (FEP_Post).

Fig. S12 shows the FEP_Post symptom correlation matrix rearranged by the 4 communities detected with community detection. Symptom community 1 included all the positive symptoms, and community 2 all of the negative symptoms.

Fig. S13 shows the three population communities that were detected, with 282, 242, and 163 patients in each community PN, Pn, and P.

Fig. S14 shows the average symptom values in the three patient clusters for the FEP_Post group. A comparison of Fig. 3c and Fig. S14 shows that the community detection method (Fig. S14) produced almost the same mean values for the symptoms of the three groups as the k-means analysis (Fig. 3c), although the numbers of patients in each group were somewhat different.

Discussion of Community detection results

The analyses with the community detection methods described separated each *population*, MEP and FEP, into three subpopulation communities. The mean values of the symptoms in these subpopulation communities were very similar to, indeed almost identical to, those shown in the main paper that were identified with k-means clustering. Moreover, this applied to both groups of patients, MEP and FEP, both when unmedicated and when medicated. Moreover, fast-mo_sgn, designed to facilitate the analysis of correlation matrices with some negative correlation (2), in practice with the data analyzed here produced essentially the same classification as fast_mo (1). An implication is that community detection is not providing something with this dataset that is conceptually different and goes beyond what k-means shows, though an advantage of the community detection approach used is that it does identify an optimal number of clusters into which to classify the data.

Interestingly, the community detection method when applied to the *symptom* correlation matrix identified in most cases four symptom communities, with the majority of the *negative* and of the *positive* symptoms in separate communities, but clustered with some of the general symptoms (Figs. S1 and S8). In both patient groups, both before and after treatment there was an '*affective*' symptom community dominated by symptoms of depression and anxiety: somatic concern G1; anxiety G2; guilt G3; tension G4 and depression G6. (Community 2 in MEP-Post; and community 3 in MEP-Pre, FEP-Pre and FEP-Post). In the FEP-Pre group, hallucinations are also included in community 3. Furthermore, in the MEP-Pre group, the other community (2, see Fig. S1) was dominated by symptoms reflecting *disorganization* of thought and behavior, together with

cognitive dysfunction: Conceptual disorganization P2; Difficulty in abstract thinking N5; Stereotyped thinking N7; Mannerisms and posturing G5; Disorientation G10; Poor attention G11; and Preoccupation G15. A similar community was identified in the FEP-Pre group (Fig. S8). Further analyses for both the MEP and FEP group prior to treatment showed that the mean of the negative symptoms was not correlated with either the affective group of symptoms, or the depression score (G6) (r in all cases ≤ 0.05, ns), nor indeed with the mean positive symptom score for P1 and P3-P7 ($r<0.1$, ns).

Comparison with bipartite classification

The community detection method used in this Supplementary Material did not include the bipartite clustering approach that utilizes both the correlations in the populations and the correlations in the symptoms being developed by Lu et al in preparation (2) in an analysis of a partly overlapping dataset. The present results provide an analysis that adopts a more usual community detection approach (1). Further, we note that the datasets are different in this study and that analysed by Lu et al (2) in that in the present study the same FEP patient set was analysed pre- and post-medication, to enable a direct comparison, and also in that the Lu et al (2) investigation had no MEP dataset, and the MEP dataset is very large.

The results obtained with the community detection methods, using fast_mo_sgn.m modified from fast_mo.m (1) with bipartite clustering based on both the population and the symptom distributions are described elsewhere (2). We note that the bipartite analysis detected only two patient communities for the FEP group, whereas the same community detection algorithm fast_mo_sgn but without bipartite analysis detected 3 communities as described here. Apart from that difference, the bipartite classification separated the populations on a similar basis, with the means between the different communities reflecting largely differences in the Negative symptoms (2). Further, the bipartite categorization sometimes found 3 patient groups for the FEP_Pre group, the Q value indicating the optimality of the solution was very similar when the bipartite method found 2 and when it found 3 groups.

Factor analysis and multidimensional scaling on the symptom correlation matrices

To provide further insight into how the different PANSS symptoms are related to each other, and how they separate from each other, factor analyses and multidimensional scaling were performed on the symptom correlation matrices. The functions available in Matlab were used.

Fig. S15 shows the results of factor analysis on the MEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms contribute to factor 1; that most of the positive symptoms contribute to factor 2; and that some of the general symptoms, most notably the affective symptoms, G1,G2,G3,G4 and G6, contribute to factor 3. In that factor 1 relates to the greatest amount of variance, this provides further evidence that variation in the negative symptoms are the major source of variation between individuals with schizophrenia. This was confirmed with principal component analysis, which showed that the first principal component accounted for 16.7% of the variance, with the second 11.4%, and the third 7.4%.

Fig. S16 shows the results of multidimensional scaling on the MEP pretreatment symptom correlation matrix. The distances in this 2D space reflect how similar each symptom is to the others, assessed across the whole population of patients (8). This shows that most of the negative symptoms are grouped close together in one part of the space; that most of the positive symptoms are in a different part of the space, apart from P2 which is placed towards the Negative symptoms; and that some of the general symptoms including the affective symptoms, are separate from the Negative and Positive symptoms, while other general symptoms are close to either negative or Positive symptoms. This Figure provides considerable insight into the relation between the different symptoms.

Fig. S17 shows the results of factor analysis on the FEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms contribute to factor 1; that most of the positive symptoms contribute to factor 2; and that some of the affective general symptoms contribute to factor 3. In that factor 1 relates to the greatest amount of variance, this provides further evidence that variation in the negative symptoms are the major source of variation between individuals with schizophrenia. This was confirmed with principal component analysis, which showed that the first principal component accounted for 18.8% of the variance, with the second 10.8%, and the third 7.4%.

Fig. S18 shows the results of multidimensional scaling on the FEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms are grouped close together in one part of the space; that most of the positive symptoms are in a different part of the space, apart from P2 which is placed towards the Negative symptoms; and that some of the general symptoms are separate from the Negative and Positive symptoms, while other general symptoms are close to either negative or Positive symptoms. This Figure provides considerable insight into the relation between the different symptoms.

In general, the distances between the symptoms are similar for the MEP and FEP populations, as illustrated in Figs. S17 and S18. Further, the MDS spaces were similar before treatment as illustrated and after treatment (not illustrated).

Comparison of MEP and FEP groups

It is evident from Fig. 5 that in the pre-treatment condition, the FEP and MEP groups had almost identical scores, and this is confirmed by the explicit comparison shown in Fig. S19. It is evident from Fig. 5 that in the post-treatment condition, the FEP and MEP groups had almost identical scores (though lower than in the pre-treatment groups), and this is confirmed by the explicit comparison shown in Fig. S20.

The negative symptoms of schizophrenia, and the effects of medication

In the results described here, medication reduced the negative symptoms of schizophrenia. This reduction might have been a direct effect of the medication on the negative symptoms, or it might have been an effect secondary to the reduction of the positive symptoms, or indeed, of some of the other symptoms. The following analyses were performed to assess this.

First, it was found that there was no correlation in the unmedicated state between the positive symptoms of schizophrenia (the mean of P1, and P3-P7) and the negative symptoms (the mean of N1-N7) before medication (FEP group r=-0.03, n=687; MEP group r=0.06, n=1880). After 6 weeks of medication, there was a correlation in the same individuals between the positive symptoms of schizophrenia and the negative symptoms (FEP group r=0.60, n=687; MEP group r=0.42, n=1880). These findings are consistent with either hypothesis: that the reduction of positive symptoms might have been a direct effect of the medication; or that the reduction of negative symptoms was related at least in part to the reduction in the positive symptoms. Further, there was a correlation between the negative symptoms when unmedicated and when medicated (FEP group r=-0.03; MEP group r=0.62). To assess whether the medication had an effect on the negative symptoms that was independent of the effect of the medication on the positive symptoms, an analysis of covariance (ANCOVA) was performed, using medication as the independent variable, the mean of the negative symptoms for each individual as the dependent variable, and the mean of the positive symptoms for each individual as the covariate. For the FEP group, there was a significant effect of the medication on the negative symptoms after controlling for the effect of the positive symptoms, $F(1,1371) = 33.6$, $p<0.001$. There was also a significant effect of the medication on the positive symptoms, $F(1,1371) = 87.7$, $p<0.001$). For the MEP group, there was a significant effect of the medication on the negative symptoms after controlling for the effect of the positive symptoms, $F(1,3757) = 121.8$, $p<0.001$. There was also a significant effect on the positive symptoms, $F(1,3757) = 200.0$, $p<0.001$). (Adding the additional covariate of the 'affective' symptoms described next made little difference to these results.) The correlation between the reduction of the negative symptoms and the reduction in the positive symptoms was 0.47 ($p=1.9\times10^{-38}$) for the FEP group and 0.44 ($p=1.7\times10^{-91}$) for the MEP group.

The correlations of the negative symptoms with a group of symptoms sometimes described as 'affective', G1-G4 and G6, were also assessed. There was no correlation in the unmedicated state between these 'affective' symptoms of schizophrenia and the negative symptoms (FEP group r=0.05; MEP group r=0.008). After 6 weeks of medication, there was a correlation in the same individuals between the 'affective' symptoms of schizophrenia and the negative symptoms (FEP group r=0.26, n=687; MEP group r=0.22). However, in an ANCOVA, using the affective symptoms as a covariate had little effect of the highly significant effect of the medication on the negative symptoms. There was a highly significant effect of the medication on the affective symptoms even when the negative symptoms were used as a covariate (MEP group $F(1,3757) = 286.0$, $p<0.001$), and interestingly this effect disappeared when the positive symptoms were instead used as the covariate. The correlation between the reduction of the negative symptoms and the reduction in the affective symptoms was 0.37 ($p=1.2\times10^{-23}$) for the FEP group and 0.21 ($p=1.1\times10^{-20}$) for the MEP group.

We also tested whether the 3 patient communities detected using all 30 symptoms showed differences in the effects of treatment. The reductions in the negative symptoms were similar in the different population subgroups, taking into account that the means for the different subpopulations were different. For the MEP group the medication reduced the mean of the negative symptoms for the PN community (n=770) from 3.86 by 1.17 ± 0.85 (mean ± standard deviation) $p=9.1\times10^{-154}$; for the Pn community (n=318) from 2.93 by 0.72 ± 0.85 $p=4.8\times10^{-36}$; and for the P community (n=792) from 2.63 by 0.69 ± 0.65 $p=2.1\times10^{-65}$. For the FEP group the medication reduced the mean of the negative symptoms for the PN community (n=264) from 3.77 by 1.15

± 0.98 (mean ± standard deviation) $p=1.5x10^{-39}$; for the Pn community (n=136) from 3.23 by 1.12 ± 0.77 $p=1.9x10^{-35}$; and for the P community (n=287) from 2.34 by 0.66 ± 0.65 $p=4.7x10^{-24}$.

To provide further information about the negative symptoms, the correlations between the different negative symptoms for the MEP group (see Fig. 1) are shown here, arranged in order N1 to N7 for the unmedicated state:

	N1	N2	N3	N4	N5	N6	N7
N1	1.00	0.73	0.50	0.61	0.32	0.47	0.31
N2	0.73	1.00	0.47	0.68	0.36	0.45	0.35
N3	0.50	0.47	1.00	0.50	0.32	0.58	0.19
N4	0.61	0.68	0.50	1.00	0.33	0.48	0.30
N5	0.32	0.36	0.32	0.33	1.00	0.40	0.50
N6	0.47	0.45	0.58	0.48	0.40	1.00	0.30
N7	0.31	0.35	0.19	0.30	0.50	0.30	1.00

To provide further evidence on whether the negative symptoms could be separated into different ‘subtypes’, the community detection algorithm was run on the correlations shown above between the negative symptoms. For the MEP group, the community detection algorithm fast_mo_sgn placed the negative symptoms N1 (blunted affect), N2 (emotional withdrawal), and N4 (Passive/apathetic social withdrawal) into one community; N3 (poor rapport) and N6 (Lack of spontaneity and flow of conversation) into a second community; and N5 (), Difficulty in abstract thinking) and N7 (Stereotyped thinking) into a third community, with a value of Q of 0.160. The bases for this categorization can be seen in the correlation matrix. This can be compared with a recent overview which divided the negative symptoms into one cluster with blunted affect and alogia, and a second cluster with anhedonia, avolition and asociality (9). (Alogia is reduction in the quantity of speech and in its spontaneous elaboration.)

The correlations between the different negative symptoms for the FEP group (see Fig. 3) are shown here, arranged in order N1 to N7 for the unmedicated state:

	N1	N2	N3	N4	N5	N6	N7
N1	1.00	0.75	0.55	0.69	0.35	0.53	0.31
N2	0.75	1.00	0.51	0.72	0.38	0.44	0.36
N3	0.55	0.51	1.00	0.55	0.36	0.61	0.25
N4	0.69	0.72	0.55	1.00	0.38	0.52	0.33
N5	0.35	0.38	0.36	0.38	1.00	0.44	0.47
N6	0.53	0.44	0.61	0.52	0.44	1.00	0.32
N7	0.31	0.36	0.25	0.33	0.47	0.32	1.00

For the FEP group, the community detection algorithm fast_mo_sgn placed the negative symptoms N1, N2, and N4 into one community; N3 and N6 into a second community; and N5 and N7 into a third community, with a value of Q of 0.143.

In terms of the effects of medication, these were similar across these three negative symptom communities. The mean of the 7 negative symptoms when unmedicated was 3.06, when medicated was 2.14, with a reduction of 0.92 ($p=4x10^{-65}$). The mean of the community 1 negative symptoms when unmedicated

was 3.25, when medicated was 2.32, with a reduction of 0.93 ($p=3x10^{-48}$). The mean of the community 2 negative symptoms when unmedicated was 3.36, when medicated was 2.22, with a reduction of 1.13 ($p=6x10^{-70}$). The mean of the community 3 negative symptoms when unmedicated was 2.49, when medicated was 1.79, with a reduction of 0.70 ($p=7x10^{-32}$).

Further, in terms of the effects of medication, the medication produced similar decreases in all the symptoms in the pre-medication identified PN, Pn, and P subpopulations of both the MEP and the FEP groups, as illustrated in Figs. S21 and S22, with the actual reduction as expected influenced by how high the score for each symptom was in the pre-medication state.

Overall, our findings indicate that some but clearly not all of the variance between patients in the response of the negative symptoms to medication can be accounted for by change in the positive symptoms but not by change in the affective symptoms. This is consistent with the possibility that the negative symptoms in our sample are in part secondary, in the sense that they arise as a reaction to positive symptoms, but it is also possible that the shared variance in treatment response reflects a shared primary effect of the medication on the mechanisms that produce the different symptoms. In any case, the effects of the medication on the negative symptoms are not largely accounted for by changes in the positive symptoms, as shown by the results of the ANCOVAs. This interpretation is consistent with the interpretation (10, 11) that the positive correlation between response of positive and negative symptoms to antipsychotics supports the hypothesis that different symptom domains in schizophrenia may depend on each other through a unified upstream pathological disease process. But at the same time, our findings show that the effects of the medication on the negative symptoms are to a considerable extent independent of the effects of the medication on the positive symptoms, so that it is likely that there are at least partly independent mechanisms for the positive and negative symptoms.

The evidence that an appreciable proportion of the variance in treatment response of negative symptoms can be accounted for by variance in response of positive symptoms suggests that the causes of negative symptoms are not homogeneous. In particular, it is consistent with the possibility that some negative symptoms are secondary to the positive symptoms. PANSS ratings do not distinguish primary from secondary negative symptoms. Nor do they attempt to distinguish transient negative symptoms for persistent negative symptoms. Thus, the findings of our study neither confirm nor refute the proposal a minor proportion of cases of schizophrenia suffer from the discrete type of schizophrenia characterized by persistent, primary negative symptoms which Kirkpatrick and colleagues have called Deficit Schizophrenia (12-14). However, the present results do show that in the unmedicated state, there is no correlation between the positive and negative symptoms across large populations of patients; and that medication introduces a correlation between the negative and the positive symptoms.

There is also discussion in the literature of whether there are subtypes of patients with different negative symptoms (that is, separate categories of patients); or whether a dimensional view is more appropriate (with continuous variation of the negative symptoms across the population) (15-17). The present results are relevant to this, for they show a continuous distribution of the negative symptoms across both the FEP and MEP populations with a unimodal not bimodal or multimodal distribution (Figs. 2 and 4). The present results thus support a dimensional interpretation of the differences in the negative symptoms between individuals with schizophrenia. Consistent with this, when clustering / community detection algorithms are forced to divide the

patients into subpopulations, the major difference between the subpopulations is in the negative symptoms, as shown here. Thus although algorithms such as k-means and community detection can separate data such as these into different subpopulations, that does not provide evidence that the subpopulations are highly distinct, and indeed the further analyses provided in Figs 2 and 4, and in the Supplementary material including factor analysis, multidimensional scaling, and analysis of the effects of the treatment in the different subpopulations, provide evidence that the main source of the difference between patients is in the mean of the negative symptoms, which has a continuous unimodal distribution, with the data this being closer to a multidimensional view compared to a discrete subtype view of schizophrenia.

Social, cultural and ethnic influences on response to treatment

As reported in the main text, we observed that negative and positive symptoms exhibited a similar response to treatment in both FEP and MEP cases, and furthermore that MEP and FEP cases exhibited a similar overall treatment response.

One issue to consider is the potentially beneficial effects of family or social support and of cultural attitudes. In an attempt to explain the better outcome of schizophrenia in non-industrialised countries reported in both the WHO International Pilot Study of Schizophrenia (18) and WHO Collaborative Study on the Determinants of Outcome of Severe Mental Disorders (19), Cooper and Sartorius (20) proposed that patients in non-industrialised countries might experience 'an interpersonal environment that is characterized by many relationships that are potentially supportive and flexible….[and] comparatively low level of pressure towards differentiation and specialization in work.' Cohen (21) has argued that even in that era, ethnographic studies did not support the hypothesis of Cooper and Sartorius. It is perhaps even more questionable that the supportive family environment proposed by Cooper and Sartorius would match that of modern-day China.

Nonetheless, Markus and Kiayama (22) presented evidence that East Asian culture promotes a view of one's self as being interconnected with members of one's community, in contrast to Western culture that encourages a view of one's self as an autonomous, independent and unique individual. More recently, Han and Humphreys (23) reviewed evidence from functional brain imaging studies indicating that this cultural difference in the view of self is associated with differences in the function of brain regions such as medial frontal cortex that support cognitive and affective processes. It is plausible that cultural differences between East Asians and Westerners might affect brain development leading to differences in the adaptability of the brain circuits implicated in negative symptoms, leading to greater responsivity of negative symptoms to treatment.

Ninety-five percent of the patients participating in our study lived with family or a partner. The participants in our study were in-patients throughout the period of the study, reducing the opportunity for family or social support to have a direct influence during that period. However, the evidence reviewed by Han and Humphreys (23) suggests that it is longstanding family and social influence that are more relevant to any modification of the brain circuits likely to be implicated in negative symptoms. Thus it is plausible, though speculative, that effects of family or social support and of cultural attitudes might have contributed to the relatively good treatment response of negative symptoms observed in our study.

It is also possible that genetic variations that influence pharmacokinetic or pharmacodynamic processes might in principle account for ethnic or racial differences in response to antipsychotics. However, the evidence is inconclusive. For example, in a review of 80 clinical studies on polymorphisms in candidate genes that might influence neurotransmitters or receptors, Kirchheiner et al (24) did not find consistent evidence of significant associations between potentially relevant genotypes and either therapeutic response or adverse drug reactions.

With regard to our finding that MEP cases show a similar response to FEP cases, contrary to findings in Western studies, one issue to consider is the possibility of differences in the proportions of patients with schizophrenia who are prescribed continuous treatment with antipsychotic medication. It would be expected that individuals who relapse while taking antipsychotic medication are less likely to respond to subsequent antipsychotic medication than those who relapse while not receiving medication. This might arise either because those relapsing while taking antipsychotics are inherently less likely to benefit from antipsychotic treatment, or because continuous exposure to antipsychotics might have resulted in reduced effectiveness, for example by altering dopamine receptor sensitivity leading to 'supersensitivity' psychosis (25). In our study, only 32.4% of the MEP cases had relapsed while receiving antipsychotic medication. Thus, the relatively good treatment response in MEP cases might be at least partially explained by only a small proportion of cases predisposed to a poor response indicated by prior relapse while taking antipsychotic medication.

.

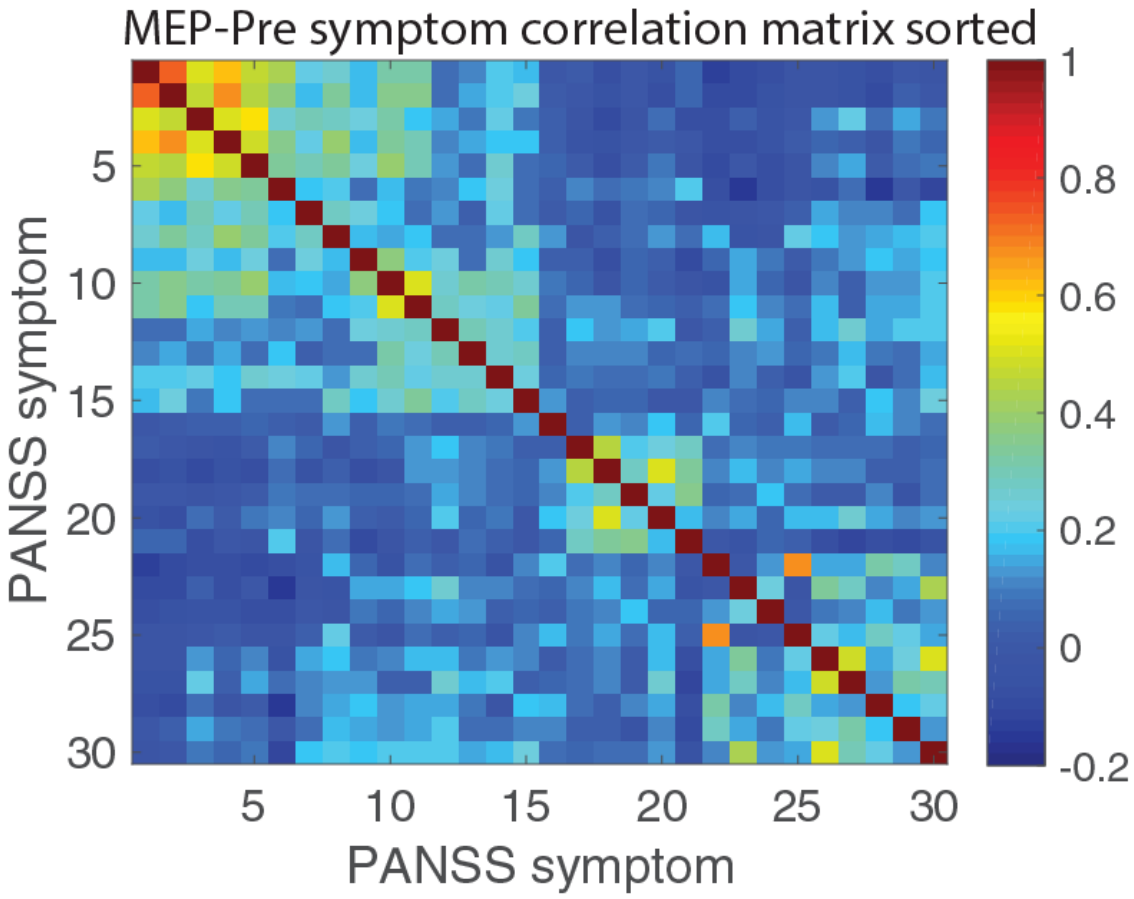


Fig. S1. MEP_Pre symptom correlation matrix, rearranged to show the 4 communities detected with the community detection algorithm. The color bar indicates the value of the Pearson correlation. The scores are arranged in this matrix according to the order below, where N1 of community 1 is shown as symptom 1 and G14 of community 4 as symptom 30. (The names provided for each community are referred to in the text.)

Community 1: N1 N2 N3 N4 N6 G7 G13 G16. (Negative)

Community 2: P2 N5 N7 G5 G10 G11 G15. (Disorganization of thought and behavior, and cognitive dysfunction).

Community 3: P3 G1 G2 G3 G4 G6. (Affective / Depressed)

Community 4: P1 P4 P5 P6 P7 G8 G9 G12 G14. (Positive)

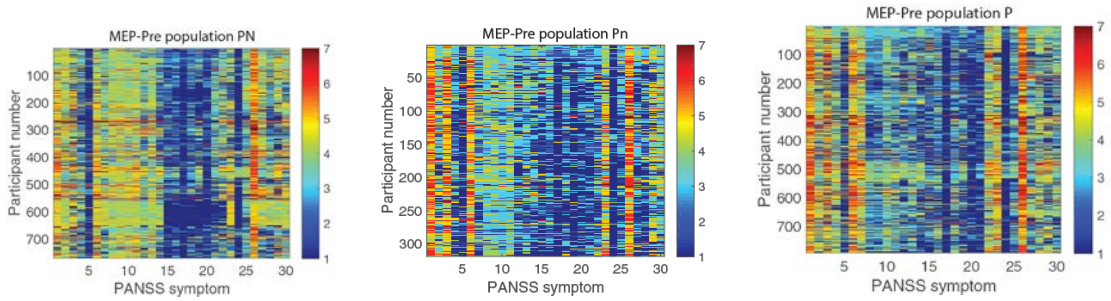


Fig. S2. Three clusters of MEP_Pre patients detected with community detection.

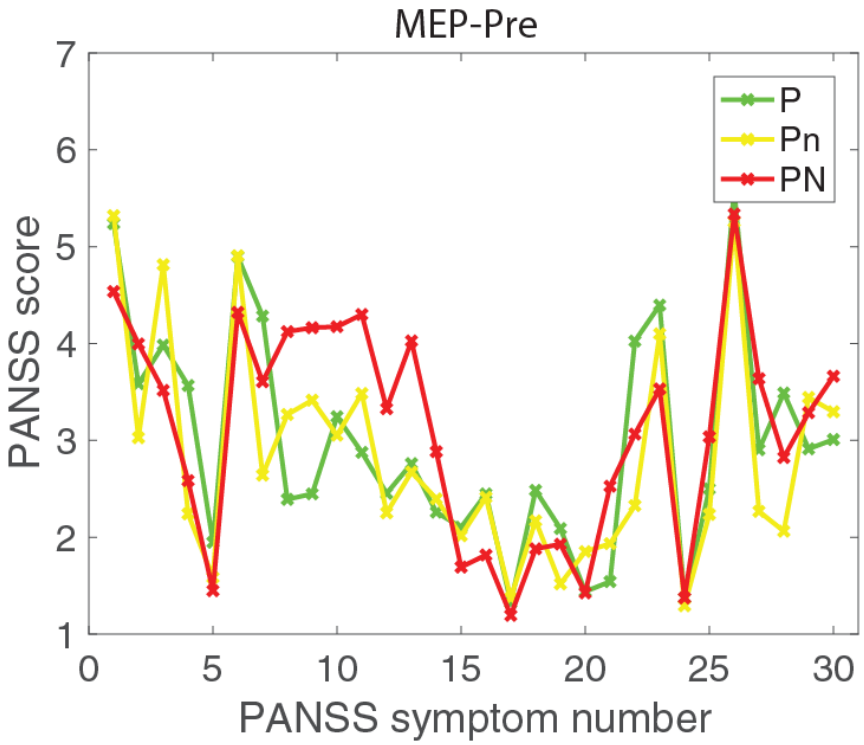


Fig. S3. MEP_Pre average symptom values in the three patient clusters detected by community detection.

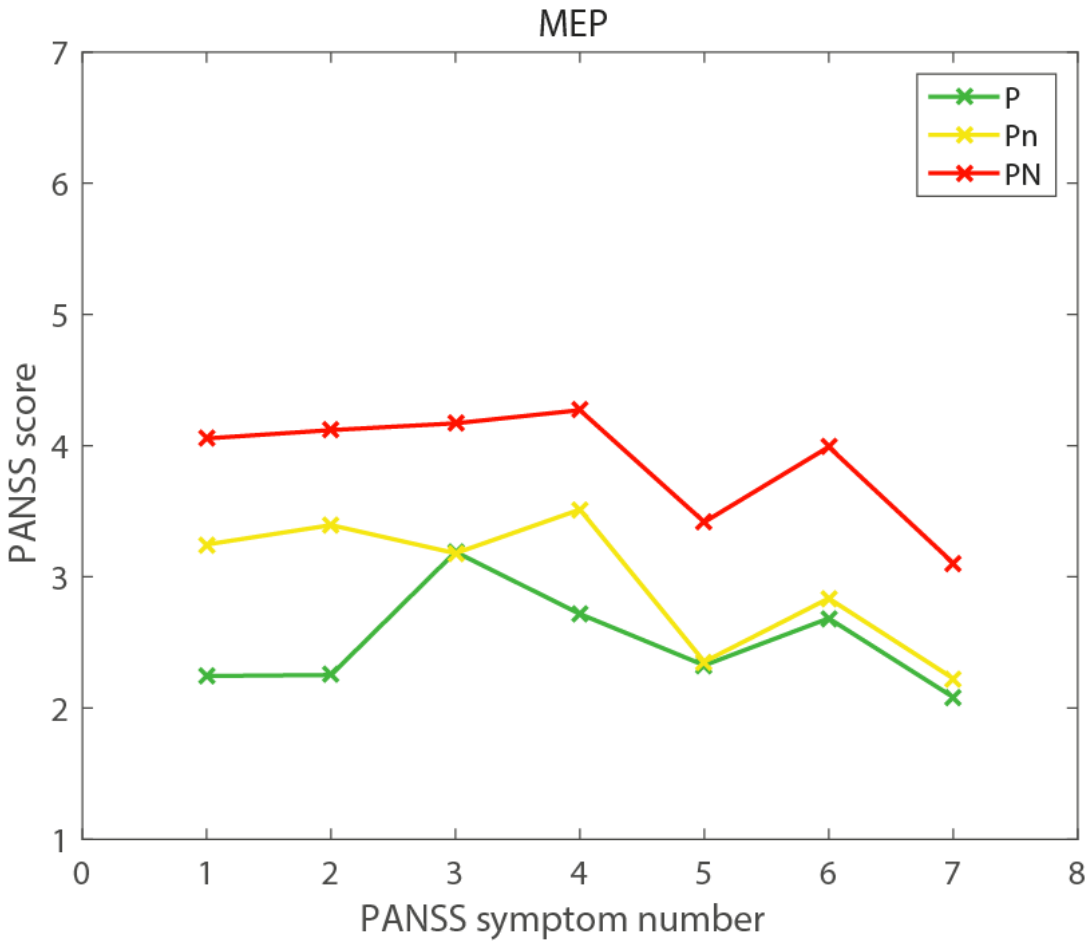


Fig. S4. MEP_Pre average symptom values in the three patient clusters that were identified by k-means with a cosine similarity measure based only on symptoms N1 – N7. The mean of the P group was 2.50, of the Pn group 2.96, and of the PN group 3.87. This shows that k-means using only the negative symptoms still categorized the patients into communities that differed in the mean value of the negative scores. (fast_mo_sgn and k-means with a correlation measure was not carried out as some patients had identical scores for the 7 negative symptoms.)

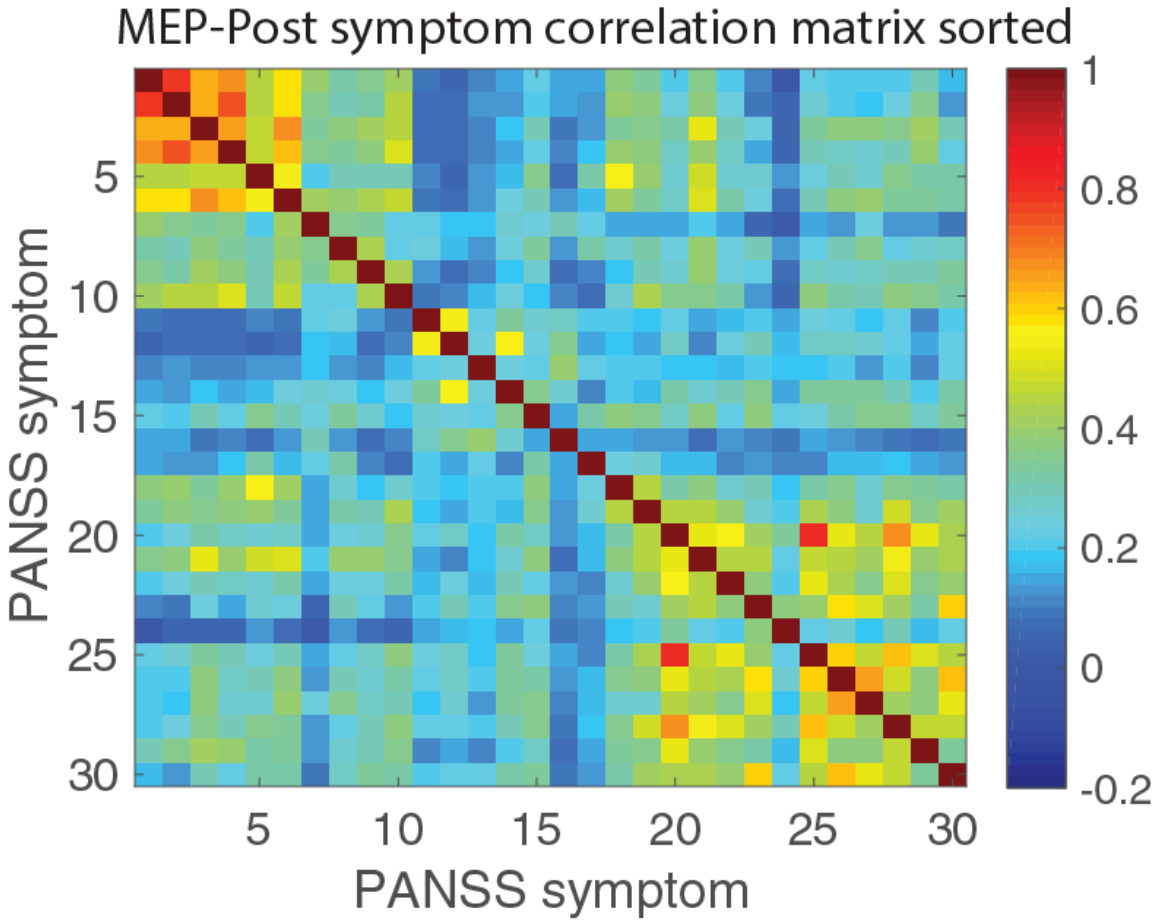


Fig. S5. MEP_Post symptom correlation matrix rearranged according to the 4 communities detected. The scores are arranged in this matrix according to the order below, where N1 is shown as symptom 1 and G14 as symptom 30.

Community 1: N1 N2 N3 N4 N6 G7 G11 G13 G16.

Community 2: G1 G2 G3 G4 G5 G6 G10.

Community 3: N7 G15.

Community 4: P1 P2 P3 P4 P5 P6 P7 G8 G9 G12 G14.

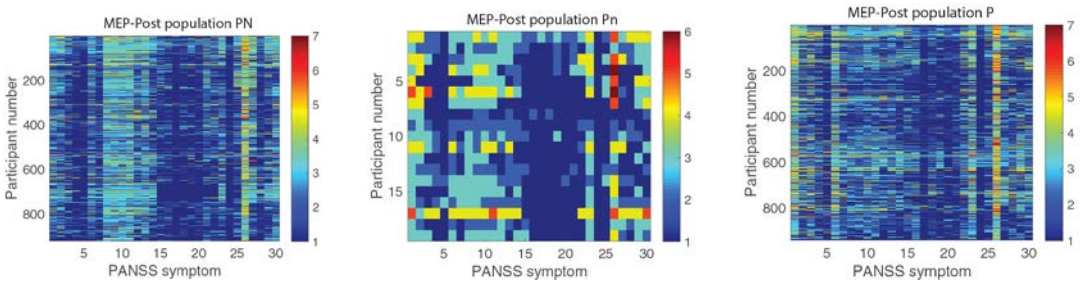


Fig. S6. Three clusters of MEP_Post patients detected by community detection.

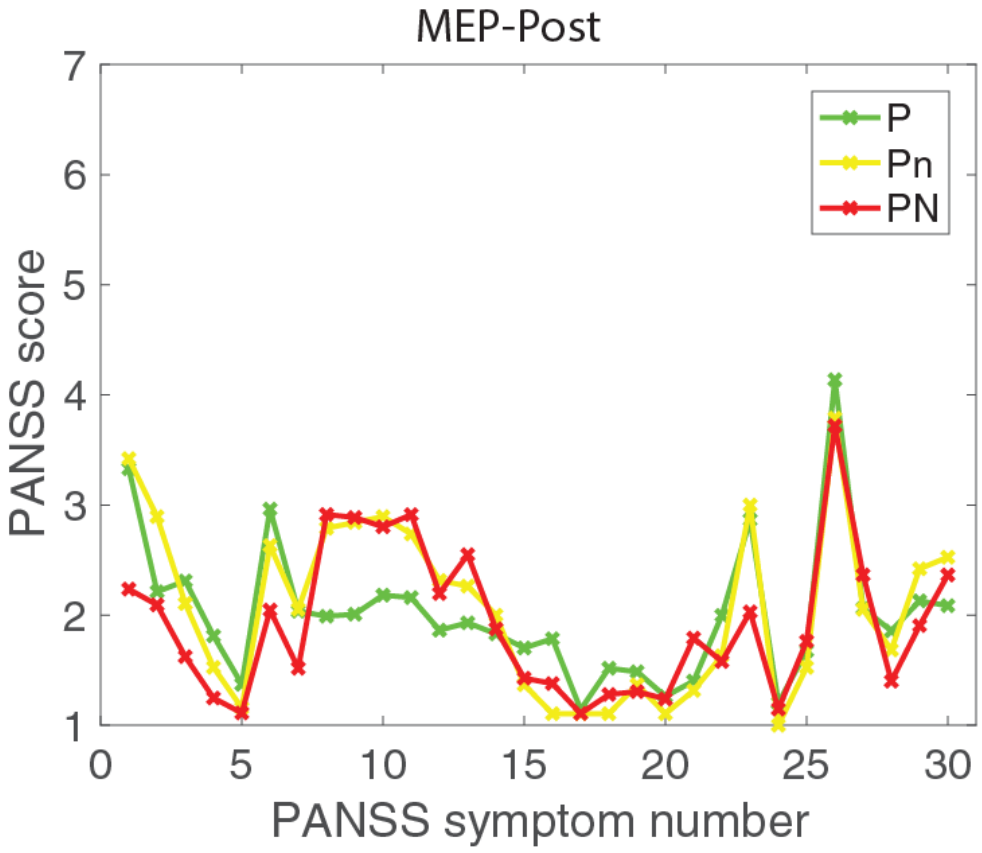


Fig. S7. MEP_Post average symptom values in the three patient clusters detected by community detection.

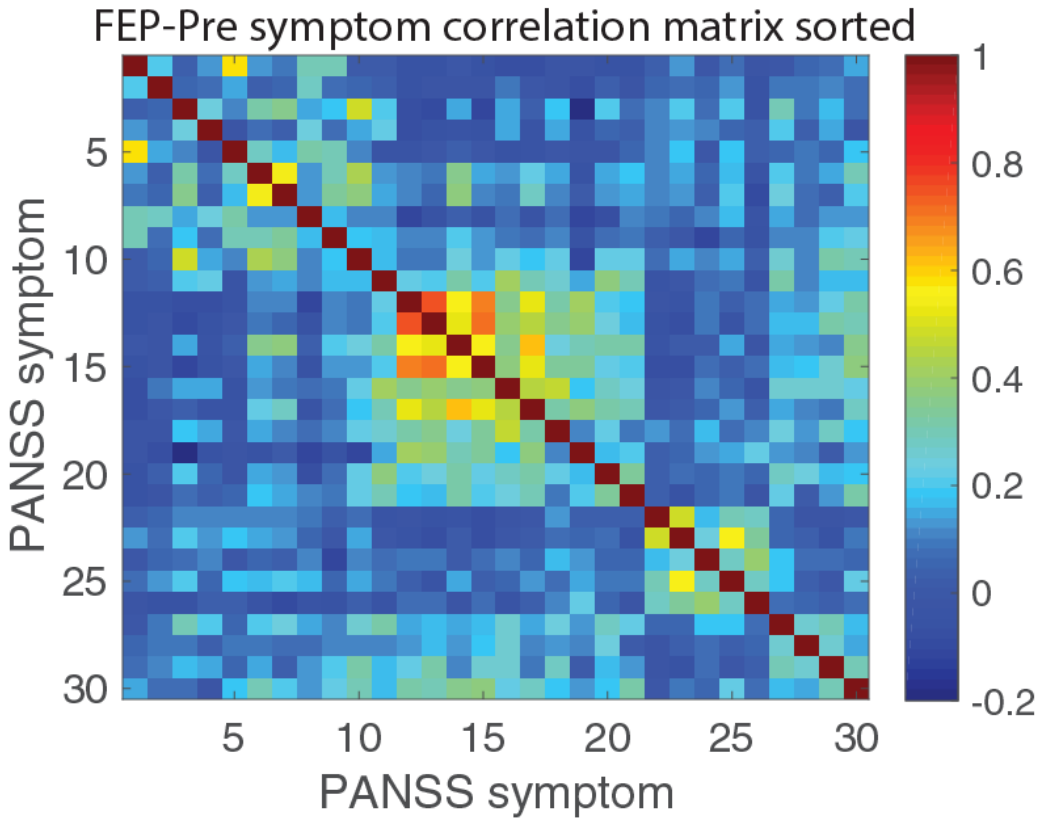


Fig. S8. FEP_Pre symptom correlation matrix rearranged by the 5 communities detected with community detection. The scores are arranged in this matrix according to the order below, where P1 of community 1 is shown as symptom 1 and G15 of community 5 is shown as symptom 30. (The names provided for each community are referred to in the text.)

Community 1: P1 P3 P4 P5 P6 P7 G8 G9 G12 G14 (Positive)

Community 2: P2 N1 N2 N3 N4 N5 N6 N7 G7 G11 G13 (Negative)

Community 3: G2 G3 G4 G6 (Affective / Depressed)

Community 4: G5 G10 G15 (Disorganization of thought and behavior, and cognitive dysfunction).

Community 5: G16

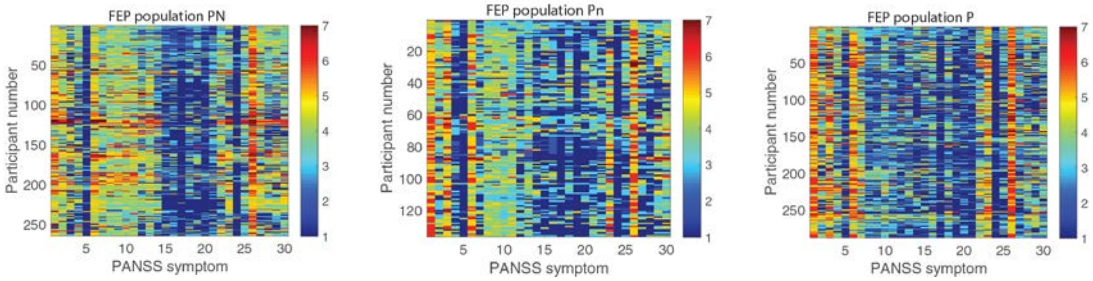


Fig. S9. Three clusters of FEP_Pre patients detected by community detection. The clusters are designated as PN, Pn and P, as they differ in the negative symptoms 8-14.

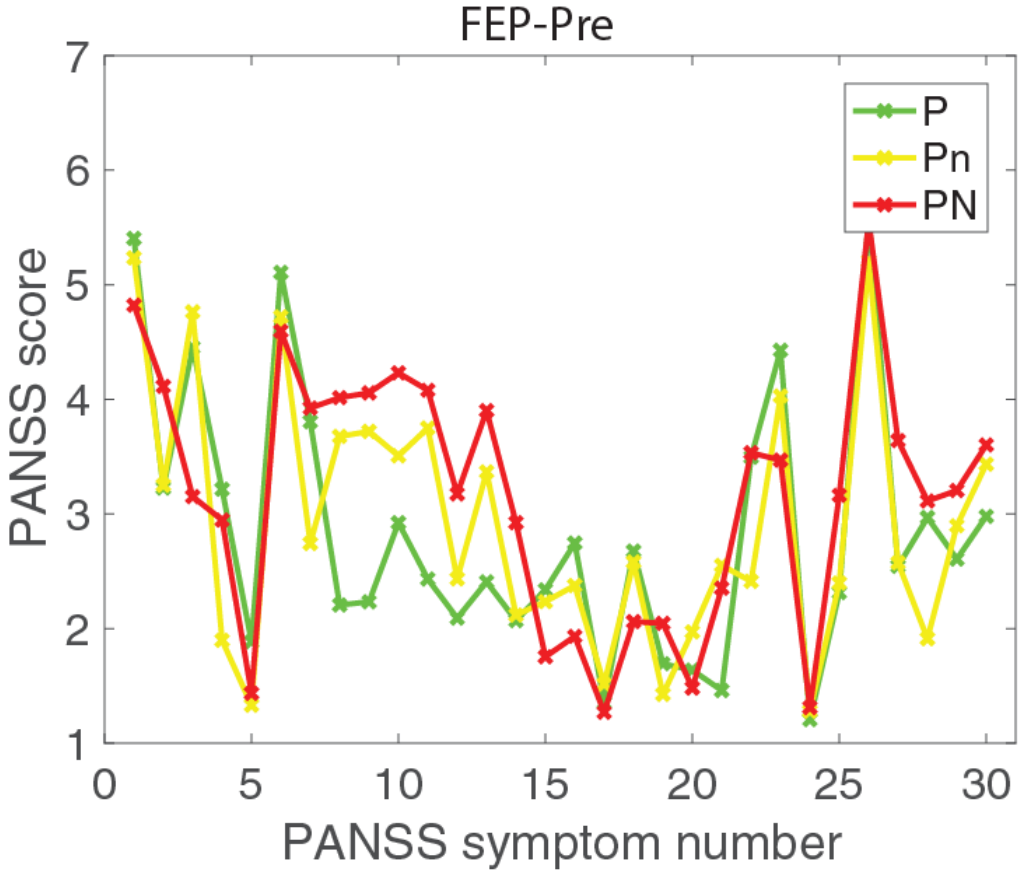


Fig. S10. FEP_Pre average symptom values in the three patient clusters detected with community detection.

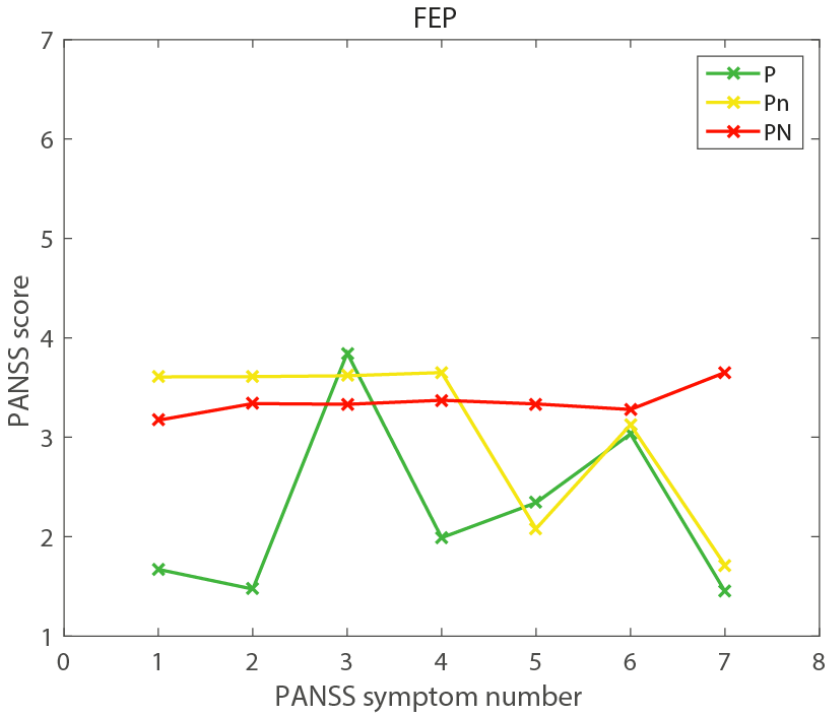


Fig. S11. FEP_Pre average symptom values in the three patient clusters that were identified by k-means with a cosine similarity measure based only on symptoms N1 – N7. The mean of the P group was 2.26, of the Pn group 3.06, and of the PN group 3.35. This shows that k-means using only the negative symptoms still categorized them into communities that differed in the mean value of the negative scores. (fast_mo_sgn and k-means with a correlation measure was not carried out as some patients had identical scores for the 7 negative symptoms.)

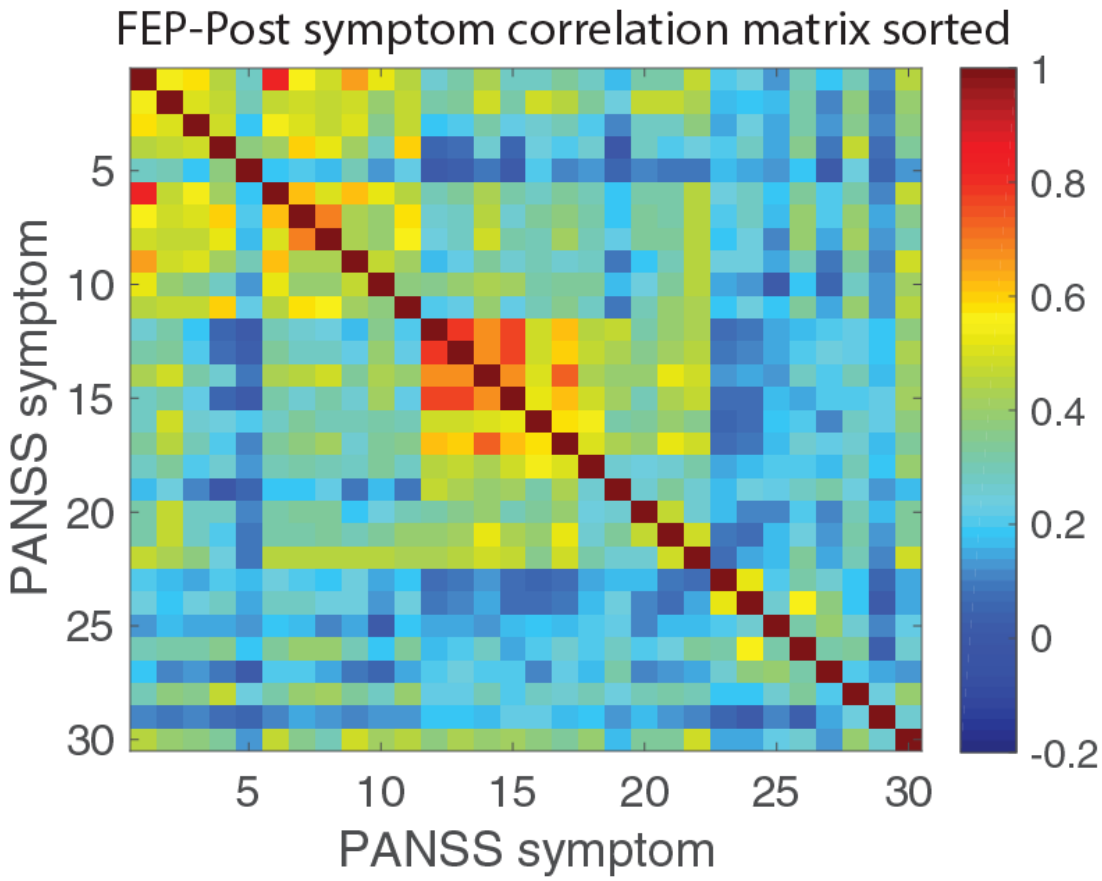


Fig. S12. FEP_Post symptom correlation matrix rearranged by the 4 communities detected with the community detection algorithm. The scores are arranged in this matrix according to the order below, where P1 is shown as symptom 1 and G15 as symptom 30.

Community 1: P1 P2 P3 P4 P5 P6 P7 G8 G9 G12 G14.

Community 2: N1 N2 N3 N4 N5 N6 N7 G7 G11 G13 G16.

Community 3: G1 G2 G3 G4 G6.

Community 4: G5 G10 G15.

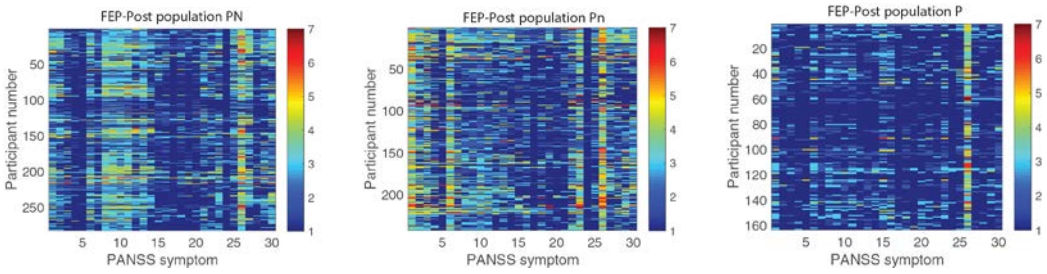


Fig. S13. Three clusters of FEP_Post patients detected with community detection.

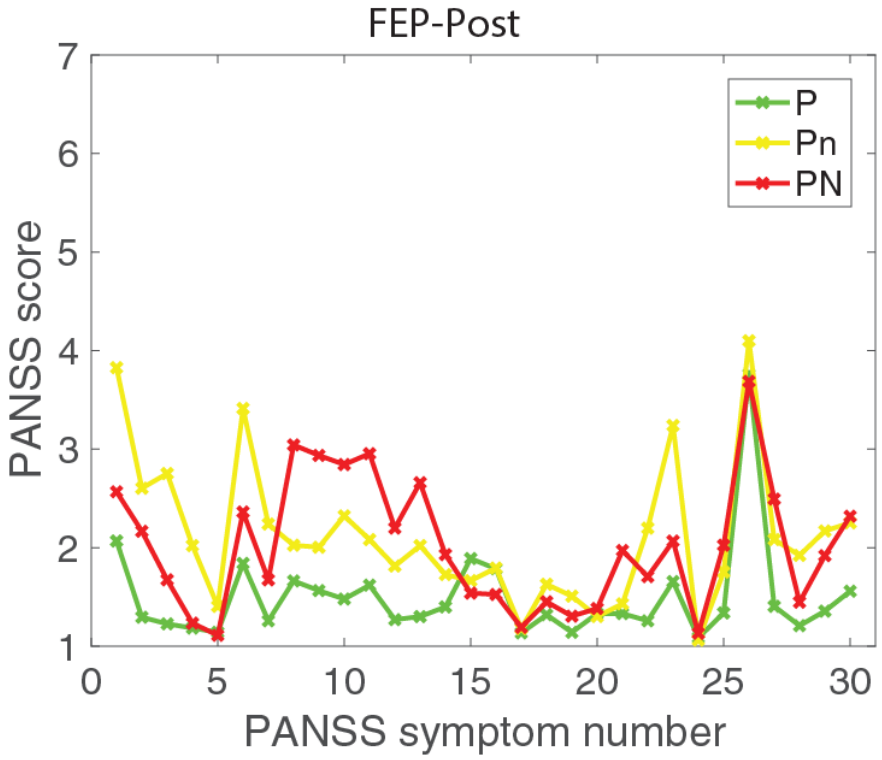


Fig. S14. FEP_Post average symptom values in the three patient clusters using community detection.

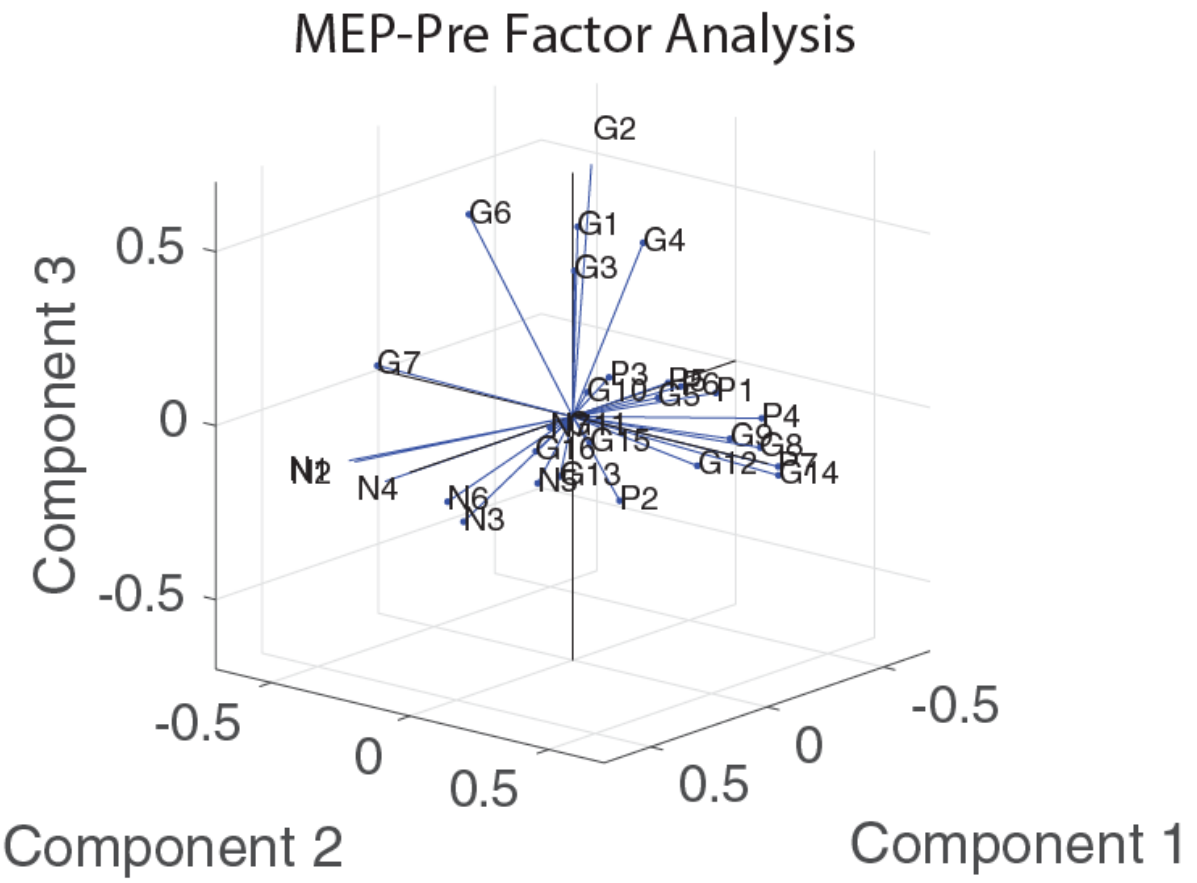


Fig. S15. Factor analysis on the MEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms are in factor 1; that most of the positive symptoms are in factor 2; and that some of the general symptoms are in factor 3.

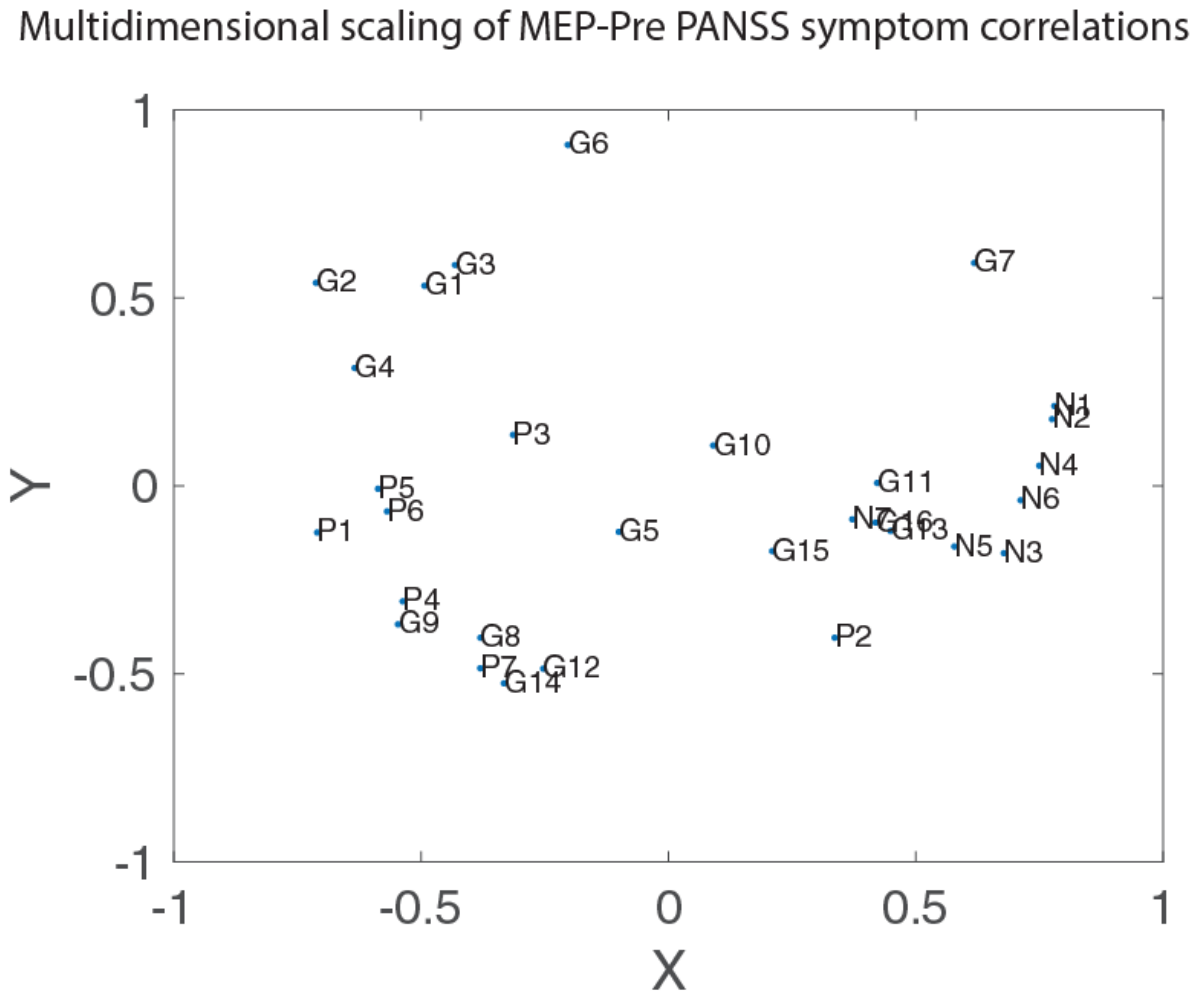


Fig. S16. Multidimensional scaling on the MEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms are grouped close together in one part of the space; that most of the positive symptoms are in a different part of the space, apart from P2 which is placed towards the Negative symptoms; and that some of the general symptoms are separate from the Negative and Positive symptoms, while other general symptoms are close to either negative or Positive symptoms. The distances in this 2D space reflect how similar each symptom is to the others, assessed across the whole population of patients.

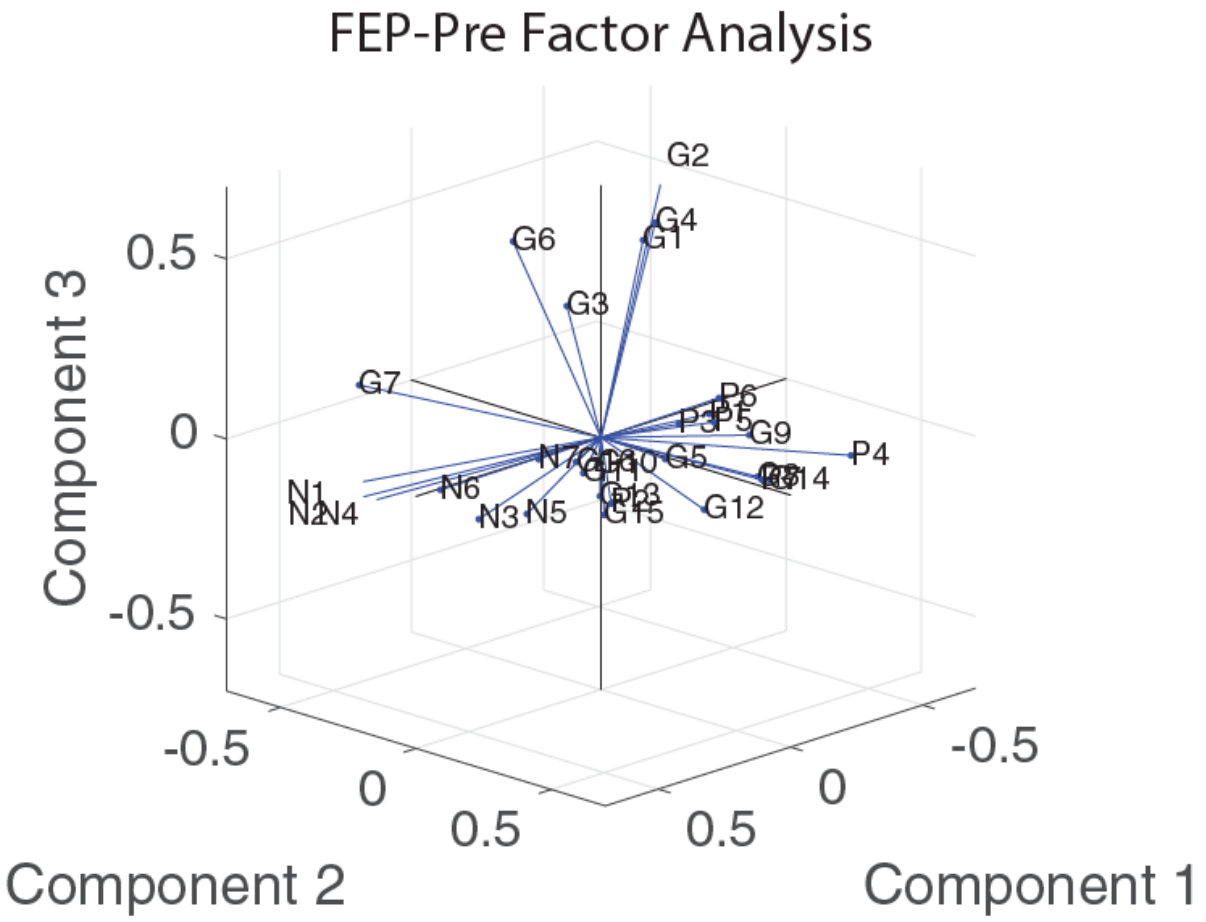


Fig. S17. Factor analysis on the FEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms are in factor 1; that most of the positive symptoms are in factor 2; and that some of the general symptoms are in factor 3.

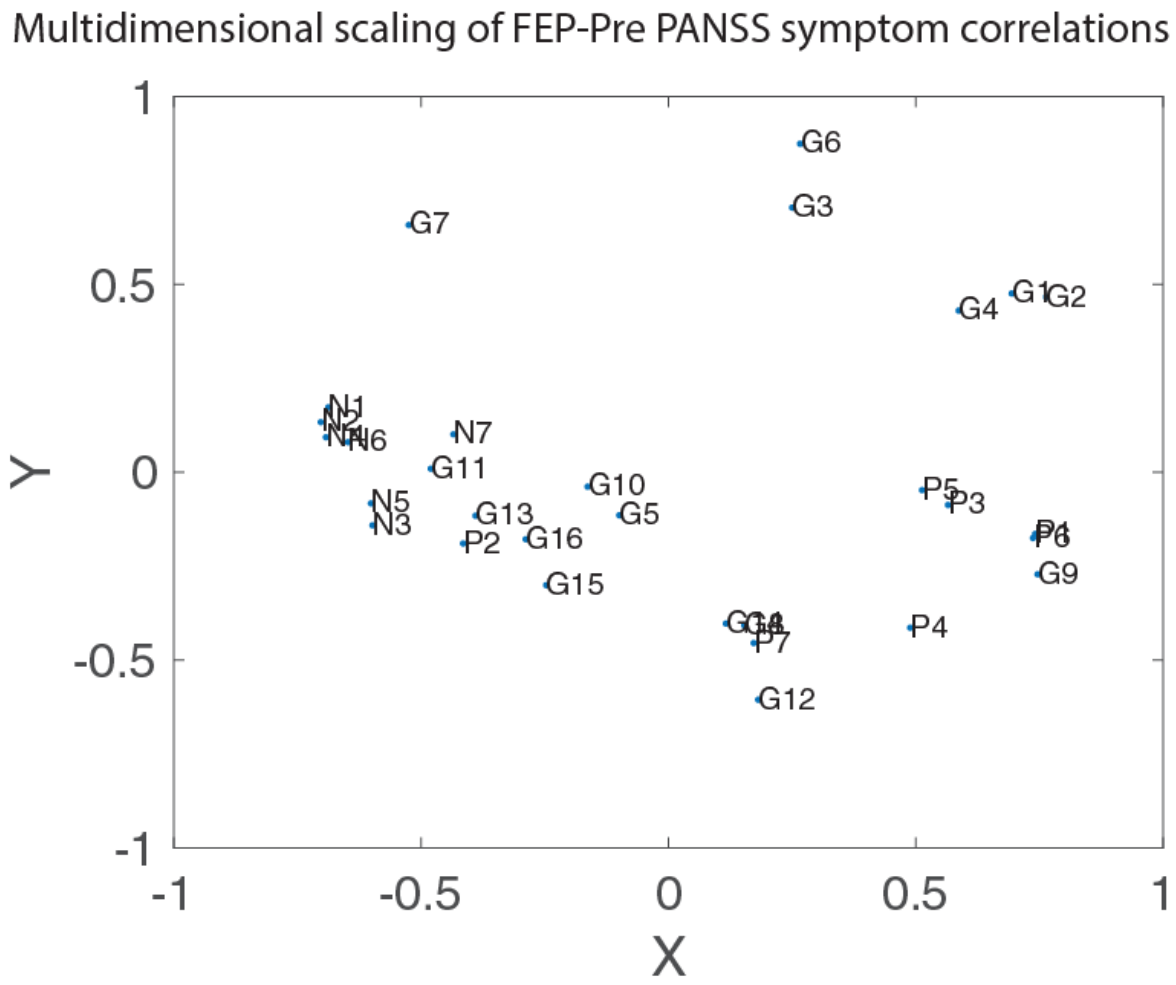
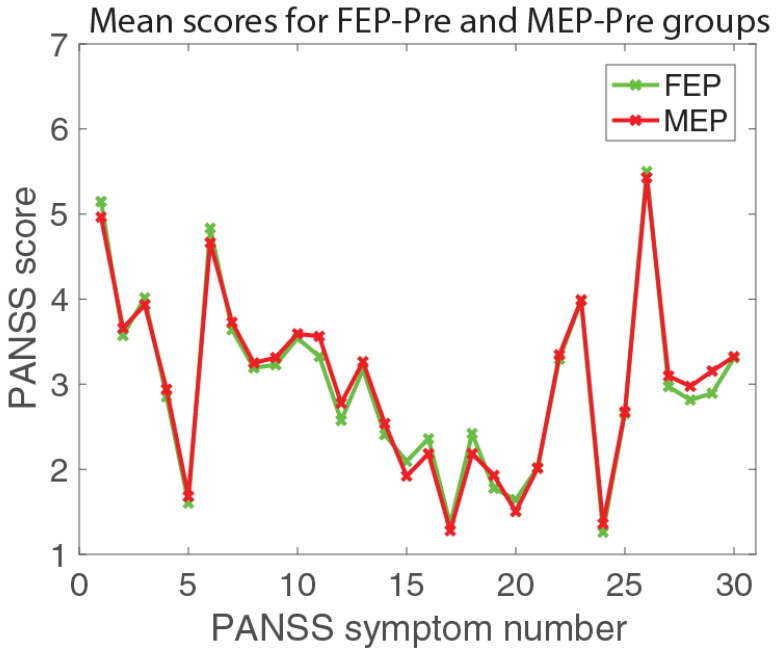


Fig. S18. Multidimensional scaling on the FEP pretreatment symptom correlation matrix. This shows that most of the negative symptoms are grouped close together in one part of the space; that most of the positive symptoms are in a different part of the space, apart from P2 which is placed towards the Negative symptoms; and that some of the general symptoms are separate from the Negative and Positive symptoms, while other general symptoms are close to either negative or Positive symptoms. The distances in this 2D space reflect how similar each symptom is to the others, assessed across the whole population of patients.



Fig, S19. Mean scores for the FEP_Pre and MEP_Pre groups. There is little difference.

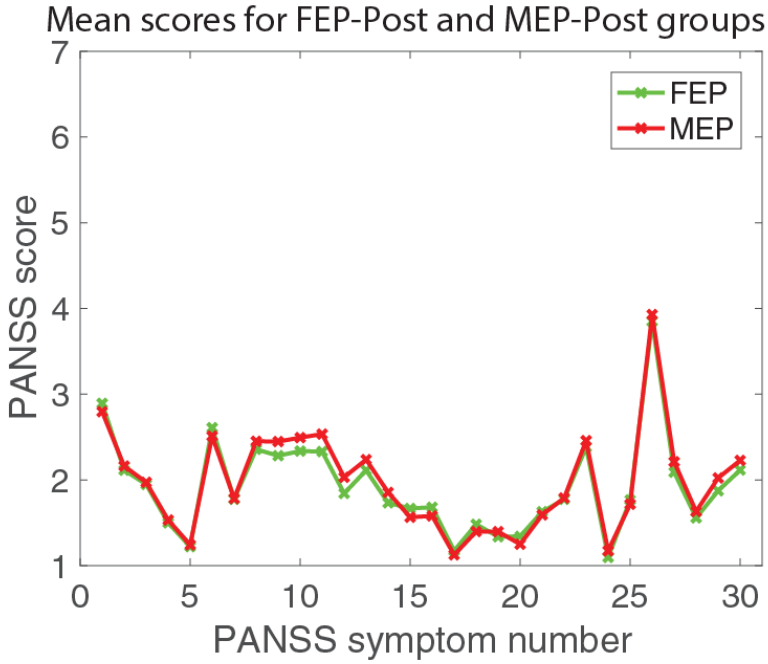


Fig. S20. Mean scores for the FEP_Post and MEP_Post groups. There is little difference between the groups, though most symptoms have decreased from the pre-treatment state.

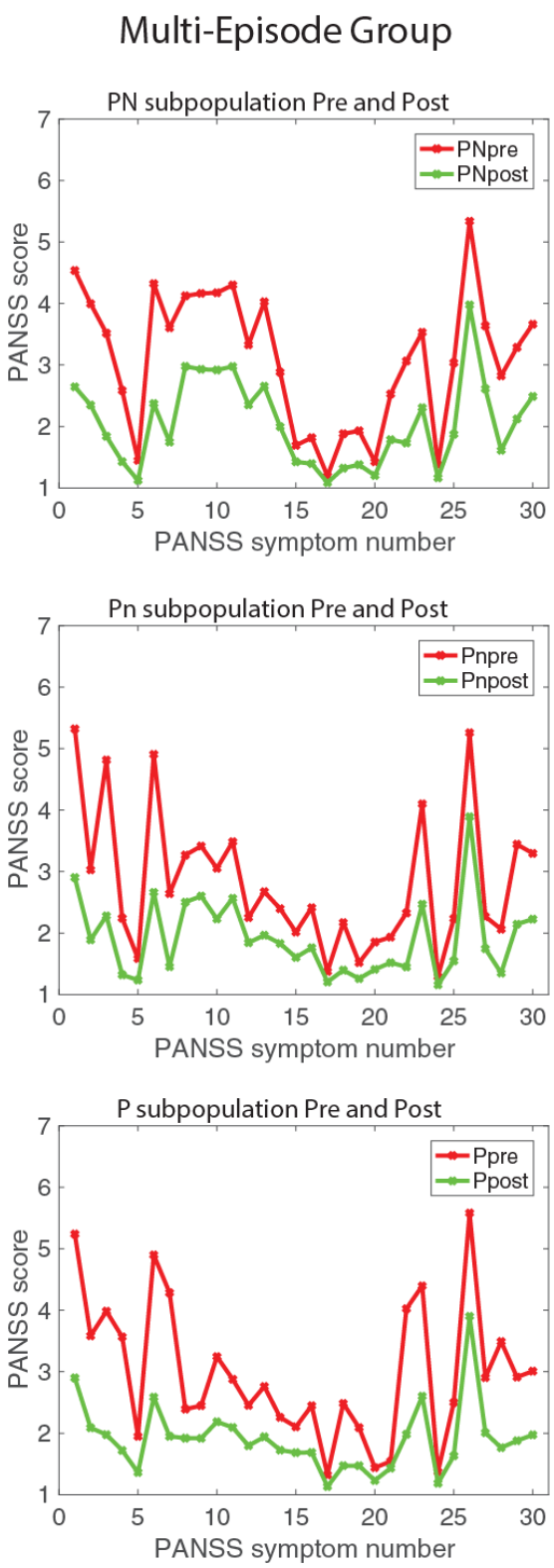


Fig. S21. The reduction in each of the PANSS symptom scores the PN, Pn, and P subpopulations identified at the pre-medication stage for the MEP group.

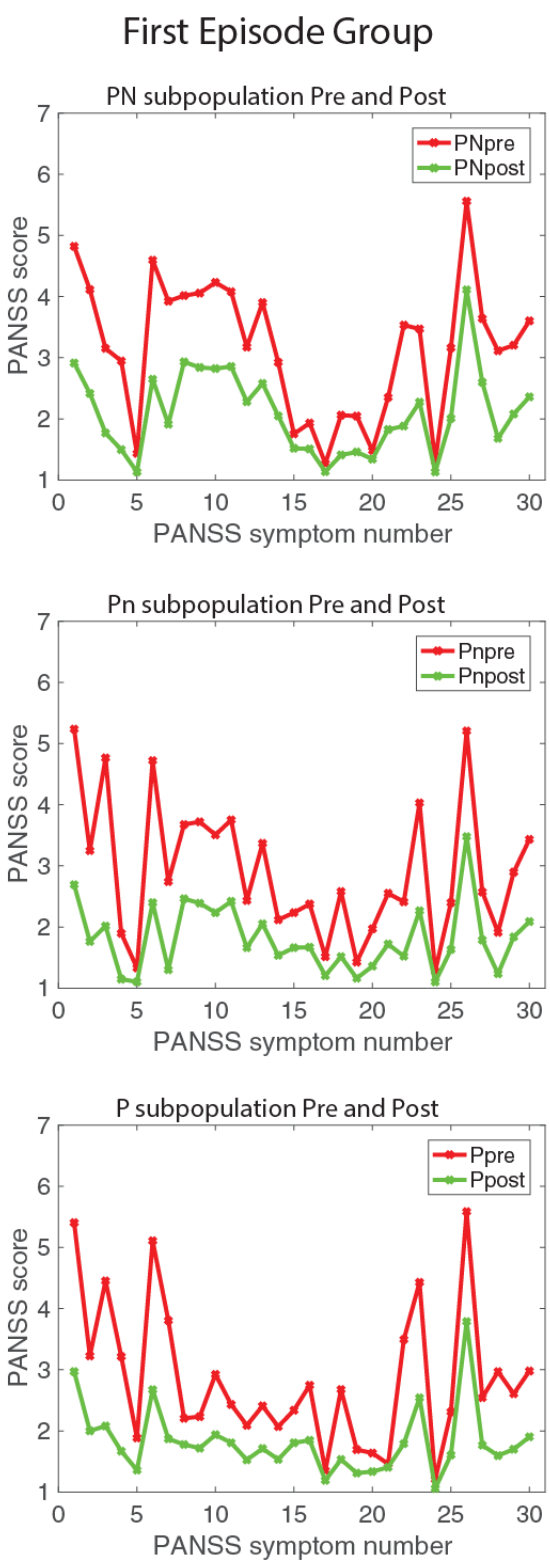


Fig. S22. The reduction in each of the PANSS symptom scores the PN, Pn, and P subpopulations identified at the pre-medication stage for the FEP group.

References

1. Le Martelot E, Hankin C. Fast multi-scale detection of relevant communities in large-scale networks. The Computer Journal. 2013; 56: 1136-50.

2. Lu W, Wan L, Rolls ET, Liddle PF, Ma L, Yan H, et al. Novel subtyping of schizophrenia predicts response to antipsychotics. In preparation. 2017.

3. Newman ME, Girvan M. Finding and evaluating community structure in networks. Physical review E. 2004; 69(2): 026113.

4. Traag VA, Bruggeman J. Community detection in networks with positive and negative links. Phys Rev E Stat Nonlin Soft Matter Phys. 2009; 80(3 Pt 2): 036115.

5. Reichardt J, Bornholdt S. Statistical mechanics of community detection. Physical Review E. 2006; 74(1): 016110.

6. DSM-IV. Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, 1984.

7. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 1987; 13(2): 261-76.

8. Schiffman SS, Reynolds ML, Young FW. Introduction to Multidimensional Scaling. Academic Press, 1981.

9. Marder SR, Galderisi S. The current conceptualization of negative symptoms in schizophrenia. World Psychiatry. 2017; 16(1): 14-24.

10. Chen L, Johnston JA, Kinon BJ, Stauffer V, Succop P, Marques TR, et al. The longitudinal interplay between negative and positive symptom trajectories in patients under antipsychotic treatment: a post hoc analysis of data from a randomized, 1-year pragmatic trial. BMC Psychiatry. 2013; 13: 320.

11. Marques TR, Arenovich T, Agid O, Sajeev G, Muthen B, Chen L, et al. The different trajectories of antipsychotic response: antipsychotics versus placebo. Psychol Med. 2011; 41(7): 1481-8.

12. Kirkpatrick B, Buchanan RW, Ross DE, Carpenter WT, Jr. A separate disease within the syndrome of schizophrenia. Arch Gen Psychiatry. 2001; 58(2): 165-71.

13. Kimhy D, Yale S, Goetz RR, McFarr LM, Malaspina D. The factorial structure of the schedule for the deficit syndrome in schizophrenia. Schizophr Bull. 2006; 32(2): 274-8.

14. Kirkpatrick B, Galderisi S. Deficit schizophrenia: an update. World Psychiatry. 2008; 7(3): 143-7.

15. Ahmed AO, Strauss GP, Buchanan RW, Kirkpatrick B, Carpenter WT. Are Negative Symptoms Dimensional or Categorical? Detection and Validation of Deficit Schizophrenia With Taxometric and Latent Variable Mixture Models. Schizophr Bull. 2015; 41(4): 879-91.

16. Strauss GP, Hong LE, Gold JM, Buchanan RW, McMahon RP, Keller WR, et al. Factor structure of the Brief Negative Symptom Scale. Schizophr Res. 2012; 142(1-3): 96-8.

17. Kirkpatrick B, Fenton WS, Carpenter WT, Jr., Marder SR. The NIMH-MATRICS consensus statement on negative symptoms. Schizophr Bull. 2006; 32(2): 214-9.

18. Leff J, Sartorius N, Jablensky A, Korten A, Ernberg G. The International Pilot Study of Schizophrenia: five-year follow-up findings. Psychol Med. 1992; 22(1): 131-45.

19. Jablensky A, Sartorius N, Ernberg G, Anker M, Korten A, Cooper JE, et al. Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization ten-country study. Psychol Med Monogr Suppl. 1992; 20: 1-97.

20. Cooper J, Sartorius N. Cultural and temporal variations in schizophrenia: a speculation on the importance of industrialization. Br J Psychiatry. 1977; 130: 50-5.

21. Cohen A. Prognosis for schizophrenia in the Third World: a reevaluation of cross-cultural research. Cult Med Psychiatry. 1992; 16(1): 53-75; discussion 7-106.

22. Markus HR, Kitayama S. Culture and the self: Implications for cognition, emotion, and motivation. Psychol Rev. 1991; 98(2): 224.

23. Han S, Humphreys G. Self-construal: a cultural framework for brain function. Current Opinion in Psychology. 2016; 8: 10-4.

24. Kirchheiner J, Nickchen K, Bauer M, Wong ML, Licinio J, Roots I, et al. Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. Mol Psychiatry. 2004; 9(5): 442-73.

25. Fallon P, Dursun S, Deakin B. Drug-induced supersensitivity psychosis revisited: characteristics of relapse in treatment-compliant patients. Ther Adv Psychopharmacol. 2012; 2(1): 13-22.

Members of Chinese Schizophrenia Collaboration Group

Dai Zhang [1], Weihua Yue [1] , Hao Yan [1], Tao Li [2], Linjiang Li [3], Chuanyue Wang[4], Fude Yang [5], Hongyan Zhang [1], Yueqin Huang [1], Xin Yu [1], Jun Yan[1], Tianmei Si[1], Wei deng [2], Xun Hu [2], Liwen Tan[3], Jiansong Zhou [3], Xin Ma [4], Qi Chen[4], Guigang Yang [5], Yunlong Tan[5], Luxian Lv [6], Hongxing Zhang [6], Yongfeng Yang [6], Keqing Li [7], Bo Du [7], Jianli Yang [8], Guangming Xu [8], Gang Zhang [9], Wenbin Ma [9], Cuicui Ma [9], Guoyang Qi [10], Zaohuo Cheng [10], Wei Wang [10], Honghui Chen [11], Xuebing Chen [11], Yunchun Chen [12], Ning Zhang [13], Rongxin Zhu [13], Jianxiong Fan [13], Congpei Zhang [14], Liying Yang[14], Zhiyong Li [14], Chuanhua Lu [15], Jisheng Tang [15], Lei Su [15], Yuping Ning [16], Yuping Liu [16], Shutao Pang [17], Guanjun Wang [17], Shenghai Wang [17], Xuanyin Huang [18], Rongke Wang [18], Huaqing Meng [19], Zhili Zou[19], Bin Hu [20], Lihua Yu [20], Tiansheng Guo [21], Guangya Liu [21], Bo Wang [22], Xueqin Yu [22], Ying Sun [23], Youguo Tan [24], Duanfang Cai [24], Ming Luo [25], Yueliang Zhang [25], Xiaoping Ge [26], Yueqing Ding [27], Jun Li [28], Haijun Wang [28], Deping Chen [29], Fuhua Zeng [29], Jun He [29], Yifei Xu [30], Guangxiang Zheng [30], Wenjun Mao [31], Wei Jian[31], Shiwu Yang [32], Chenglin Li [32].

[1] Institute of Mental Health, the Sixth Hospital, Peking University, Beijing 100191, China;

[2] West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China;

[3] Beijing Anding Hospital, Capital Medical University, Beijing 100088, China;

[4] The Second Xiangya Hospital of Central South University, Changsha 410011, China;

[5] Beijing HuiLongGuan Hospital, Peking University, Beijing 100096, China;

[6] The Second Affiliated Hospital of Xinxiang Medical University, Xinxiang, Henan, 453002, China;

[7] Hebei Mental Health Center, Baoding, Hebei 071000, China;

[8] Tianjin Anding Hospital, Tianjin 300222, China;

[9] Jinzhou Kangning Hospital, Jinzhou, Liaoning 121013, China;

[10] Wuxi Mental Health Center, Nanjing Medical University, Wuxi, 214151, China;

[11] Wuhan Mental Health Center, Wuhan, Hubei 430022, China;

[12] The First Affiliated Hospital of the Fourth Military Medical University, Xi'an 710032, China;

[13] Nanjing Brain Hospital, Nanjing, Jiangsu 210000, China;

[14] Harbin First Specialized Hospital, Harbin, Heilongjiang 150056, China;

[15] Shandong Mental Health Center, Jinan, Shandong 250014, China;

[16] Guangzhou Mental Hospital, Guangzhou, Guangdong 510370, China;

[17] Qingdao Mental Health Center, Qingdao, Shandong 266034, China;

[18] The Third People's Hospital of Mianyang City, Mianyang, Sichuan 621000, China;

[19] The First Affiliated Hospital of Chongqing Medical University, Chongqing 400042, China;

[20] Jiangxi Mental Hospital, Nanchang, Jiangxi 330029, China;

[21] Hunan Brain Hospital, Changsha, Hunan 410007, China;

[22] Chongqing Mental Health Center, Chongqing 401147, China;

[23] Liaoning Provincial Mental Health Center, Tieling, Liaoning, 112300, China;

[24] Zigong Mental Health Center, Zigong, Sichuan 643020, China;

[25] Third People’s Hospital, Panzhihua, Sichuan, 617000, China;

[26] Changsha Psychiatry Hospital, Changsha, Hunan, 410018, China;

[27] The Fifth People's Hospital of Jiujiang, Jiujiang, Jiangxi 332001, China;

[28] Veterans' Hospital of Sichuan Province, Chengdu, Sichuan, 611236, China;

[29] Ziyang Psychiatric Hospital, Ziyang, Sichuan, 641300, China;

[30] Sixth Hospital of Changchun, Changchun, Liaoning, 130052, China;

[31] Chengdu Mental Health Center, Fourth People’s Hospital, Chengdu 610036, China;

[32] Chengdu Dekang Hospital, Chengdu, Sichuan 610000, China

Individual differences in schizophrenia

Edmund T. Rolls, Wenlian Lu, Lin Wan, Hao Yan, Chuanyue Wang, Fude Yang, Yunlong Tan, Lingjiang Li, Chinese Schizophrenia Collaboration Group, Hao Yu, Peter F. Liddle, Lena Palaniyappan, Dai Zhang, Weihua Yue and Jianfeng Feng

Br J Psychiatry Open 2017, 3:265-273.
Access the most recent version at DOI: 10.1192/bjpo.bp.117.005058